

Query Match		70.8%;	Score 3595.8;	DB 6;	Length 3715;		
Best Local Similarity		98.6%;	Pred. No. 0;				
Matches 3685;		Conservative	0;	Mismatches	7;	Indels	45; Gaps 4;
Qy	1335	tgacgagagagaaacctctgttttccccacactctctctccacccctccctcctgcctcc	1394				
Db	4	TTCCGGGGAGAGAACCTCTGTCTTCCCCACTCTCTCCACCTCTCTCTCCCTTCCC	63				
Qy	1395	caccccgagtcgagcagagatcaaaagatgaaaggcagtcagggtcttcagtagca	1454				
Db	64	CACCCCGAGTGGGAG-CAGAGATCAAAAGATGAAAGGCAGTCAGTCTTTCAGTAGCCA	122				
Qy	1455	aaaaaacaacaacaacaacaagcgaataataaagaaagaaataaactcagtt	1514				
Db	123	AAAAACAAAAACAACAAAAACAAGGCGAAATAAAAAGAAAGATAATAACTCAGTT	182				
Qy	1515	cttattgcacctacttcagtgacactgaatttgaaagtgagagatttggtttttc	1574				
Db	183	CTTATTTGCACCTACTTTCAGTGGACACTGAATTTGGNAGGTGGAGATTTTGTGTTTTTC	242				
Qy	1575	ttttaagatctgggcatcttttgaatatcccttcacgtatttaagagacagactgagc	1634				
Db	243	TTTTAAGATCTGGCATCTTTTGAATCACTACCTTCAAGTATTAAAGACACAGACTGTGAGC	302				
Qy	1635	ctacgagcgagatcttgtccacgtgtctctctctgcacagactttgagctgtca	1694				
Db	303	CTAGCAGGGCAGATCTTGTCCACGtGTGTCTTCTGTGCACGAGACTTTGAGGCtGTCA	362				
Qy	1695	gagcgtttttggtgtgtctcccgcaagtcttctctctctgagcttcccgagtg99	1754				
Db	363	GAGCGCTTTTGTGCTGTCTCCGCAAGTTTCTTCTCTGGAGCTTCCCGCAGGTGGG	422				
Qy	1755	cagctagctcagcagactacgcatactacagcctgttgaactcttctagcaagagaa	1814				
Db	423	CAGCTAGCTGCAGCGAGTACCGCATCATCACGCTGTTCGAATCTTCTTGAGCAAGAGAA	482				
Qy	1815	ggcgagcggggttaaggaaagtagtggaagattcaagcaagctcaaggtggaagtga	1874				
Db	483	CGGAGCGGGGTAAAGGAAGTAGTAGTGGAAGATTCAGCCAAAGTCAAGGATGAAGTGCA	542				
Qy	1875	gttagggctgggaagggtctacctcgccgcccgttccaaagacttaocgagagctttcca	1934				
Db	543	GTTAGGGCTGGAAAGGTCTACCTCTCGCCGCGCGTCCAAAGACCTACCGAGAGCTTTCCA	602				
Qy	1935	gaatctttccagagctgagcaagtagatccagaacccggcccaagcaccacaagcc	1994				
Db	603	GAATCTGTTCAGAGCGTGGCGAAGTGATCCAGAACCCGGGCCCAAGCACCACAGAGGC	662				
Qy	1995	cgcgagcgagcacctcccgcgccagtttctgtctgtctgcagcagcagcagca	2054				
Db	663	CGCGAGCGCAGACCTCCCGGCGCCAGTTTGTGCTGCTGCT-----	701				
Qy	2055	gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	2114				
Db	702	---GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	758				
Qy	2115	tagcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	2174				
Db	759	TAGCCCCAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	818				
Qy	2175	aggcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	2234				
Db	819	AGGCCCAACAGGCTACCTGTGCTTGGATGAGAAACAGCAACCTTTCACAGCGCGAGTCGGC	878				
Qy	2235	cctggagtgcaaccccgagagaggttgcgtcccaagaccttgagcccgctggccccag	2294				
Db	879	CCTGGAGTGCCACCCGAGAGAGGTTGCGTCCACAGAGCCCTGGAGCCCGCTGGCCGCCAG	938				
Qy	2295	caaggggctgcgcagcagctgcagcacctcccgacagagatgactcagctgcctccatc	2354				
Db	939	CNAGGGGTGCGGACGAGCTGCCAGACCTCCGGAGGAGGATGACTCAGCTGCCCCATC	998				

Qy	2355	cacgttgccctgtggtggcccccactttcccgggttaagcagctgtccgtgaccttaa	2414
Db	999	CACGTTGCTCTGCTGGGCCCCACTTTCCCGGCTTAAGCAGTGTCTCGCTGACCTTAA	1058
Qy	2415	agacatcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	2474
Db	1059	AGACATCTGAGCAGGCCAGCAGCATGCAACTCTTTAGCAACAGCAGCAGAGAGCAGT	1118
Qy	2475	atccgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	2534
Db	1119	ATCCGAAGGCGAGCAGCAGCGGAGAGGAGGCGCTTCGGGGGCTCCACTTCTCTCAA	1178
Qy	2535	ggacaattacttaggggcactctcgaccattctgacaacgcccgaagagttgtgtgaagg	2594
Db	1179	CGACAAATTACTTAGGGGGCACTTCGACCATTTTCACAAACSCCAAGGAGTTGTGTAAGG	1238
Qy	2595	agtgctgggtgtccatgggacctgggtgtggagcgtctggagcagcagcagcagcagc	2654
Db	1239	AGTGTGCGTGTCCATGGGCGCTGGGTGTGGAGGCGTTGGAGCATCTGATCCAGGGGAACA	1298
Qy	2655	gcttcgggggattgcatgtacgcccacttttgaggattccacccgctgtgctccac	2714
Db	1299	GCTTCGGGGGATTCATGTACGCCCCACTTTTGGAGTTCCACCGCTGTGCTGCCAC	1358
Qy	2715	tccttgtccccattggccgaatgcaaggttctctctadacgacagcagcagcagcag	2774
Db	1359	TCCTTGTGCCCATTTGGCCGAATGCAAGGTTCTCTGCTAGACGACGCGCAGGCAAG	1418
Qy	2775	caactaagatactgctgagttatcccccttccaagggaggttaacaccaaagggctagaagg	2834
Db	1419	CACCTGAAGATACGTGTGATTTCCCTTTCAAGGGAGGTTTACACCAAGGGCTTAGAAG	1478
Qy	2835	cgagagcctaggtctctctggcagcgtgcagcagcagcagcagcagcagcagcagc	2894
Db	1479	CGAGAGCCTTAGGCTGCTTGGCAGGCGCTGCAGCAGGAGCTCCGGGACACTTGAATGCC	1538
Qy	2895	gtctacctgtctctctacaaagtcgggagcactggaagcagcagcagcagcagcagc	2954
Db	1539	GTCTACCTGTCTCTCTACAACTCCGGAGCACTGGACGAGGAGTGGTGTACCAAGAGTCG	1598
Qy	2955	cgactactaacaatttccactggtctgcccagcagcagcagcagcagcagcagcagc	3014
Db	1599	CGACTACTACAACTTTTCCACTGTGCTGTGGCGGACCGCGCCCTCCCGCTCCGCCA	1658
Qy	3015	tcaccaagctgcatacaagctggagaaacccgctggactacgagcagcagcagcagcagc	3074
Db	1659	TCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCTGGCGGTGC	1718
Qy	3075	ggcgcgcagtgctccgctatg9gggacctggcagcagcagcagcagcagcagcagcagc	3134
Db	1719	GGCGGCGCAGTGTCCGCTATGGGGACCTGGCGAGCCTGATGGCGGGGTGGTGGCGGCGG	1778
Qy	3135	cggttctgggtacccctcagcccgcttctctacctctggcacactcttctcagagcga	3194
Db	1779	CGGTCTTGGGTCAACCTTCAGCCCGCTTCTCTATCTTGGCACACTCTCTTCACAGCGCA	1838
Qy	3195	agaagccagtgattgtgaccctgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	3245
Db	1839	AGAAAGCCAGTTGTATGGACCTGTGTTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG	1898
Qy	3246	cggc	3305
Db	1899	CG	1958
Qy	3306	agcccccctacgctacactcggcccccctcaggggctg9gggcccaggaaagcagcttcc	3365
Db	1959	AGCCCCCTACGCTACACTTCGCCCCCTCAGGGGCTGGCGGCGCAGAAAGGACTTCAC	2018
Qy	3366	cgcaactgattgtgtgtacctgtggcagcagcagcagcagcagcagcagcagcagcagc	3425
Db	2019	CGCACCTGATGTGTGATACCTTGGCGGATGTTGAGAGAGTGGCTTATCCCGAGTCCAC	2078
Qy	3426	ttgtgtcaaaagcgaataaggcccttgatgatatgactactccgacaccttacggggacat	3485

```
|||||
Db 2079 TTGTCTCAAAAGCGAAATGGCCCTCGATGGATAGCTACTCGGACCTTACGGGCACAT 2138
QY 3486 qcgittgagactgcagggaacatgttttgccattgactatactttccacccagaa 3545
Db 2139 GCGTTTGGAGACTGCCAGGACCATGTTTTGCCCATTTGACTATTACTTTCCACCCAGAA 2198
QY 3546 gacctgcctgatctgtgagatgaagctcttgggtgcactatggagctctcacatgtg 3605
Db 2199 GACCTGCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACATATGGAGCTCTCACATGTGG 2258
QY 3606 aagctgaaggttcttctcaaaagagccgctgaagggaacagaagtaacctgtgcccag 3665
Db 2259 AAGCTGAAGGTCTTCTTCAAAAGAGCGCTGAAGGGAACAGAACTACTGTGCGCCAG 2318
QY 3666 cagaaatgattgcactatgataaattcgaagaaataattgccatcttgtcgtctcg 3725
Db 2319 CAGAAATGATTGCACTATTGATAAATTCGGAAGAAATAATGTCATCTTGTGCTCTCG 2378
QY 3726 gaaatgttatgaagcagggatgactctggagccggagagctgaagaaacttggtaact 3785
Db 2379 GAAATGTTATGAAGCAGGGATGACTCTGGAGCCCGAAGCTGAAGAACTTGGTAATCT 2438
QY 3786 gaaactcaggaaggaaggaggtcttcagaccacacacagccctcagggagacaccca 3845
Db 2439 GAAACTACAGGAGGAAGAGAGGCTTCCAGCACACCACCCACTGAGGAGACAAACCA 2498
QY 3846 gaagctgacagtctcacatgaaggctatgaatgcagaccatcttctgaatgtct 3905
Db 2499 GAAGCTGACAGTGTACACATTTGAAGGCTATGAATGTCAGGCCATCTTTCTGAATGCT 2558
QY 3906 ggaagccattgagccagggtgtagtgtgtgtgacagacacacacacagccagactcct 3965
Db 2559 GGAAGCCATTGACCGAGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2618
QY 3966 tgcagcttgcctctctagcctcaatgaactggagagagacagcttgtcacagtggtcaa 4025
Db 2619 TGCAGGCTTGTCTCTTAGGCTCAATGAACCTGGGAGAGACAGCTGTGTACAGTGTGTC 2678
QY 4026 gtgggccaagccttgccctgcgcgaacttcaactcaactgagcagaccagatggctgcat 4085
Db 2679 GTGGGCCAAGCCCTGCCCTGCTCCGCAACTTACAGTGGAGACAGCTGTGTACAGTGTGTC 2738
QY 4086 tcagtactcctggatgggctcatggtgtgtgccatgggctggcagctcttcacaaagt 4145
Db 2739 TCAGTACTCTCGATGGGCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2798
QY 4146 caactccagatgctacttgcctcctgatctgtgttttcaatgagtagcagcagcaca 4205
Db 2799 CAACTCCAGGATGCTACTTCCGCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAA 2858
QY 4206 gtcccgatgtacagccagtgctgcgaatgagggcacctctcgaaggtttgagtgct 4265
Db 2859 GTCCCGATGTACAGCCAGTGTGCGAATGAGGACACCTCTCAAGAGTTGGATGGCT 2918
QY 4266 ccaaatcaccocccaggaattcctgtgcatagaagcaactgctactcttcagcatttcc 4325
Db 2919 CCAATACCCCCAGGAATTCCTGTGATGAAGCACTGTCTACTTTCAGCATTTATCC 2978
QY 4326 agtggatggctgaaaaatcaaaaattcttggatgaacttcgaatgaactacacagga 4385
Db 2979 AGTGGATGGCTGAAAAATCAAAAATTCCTTTGATGAACCTCGAATGAACATACATCAAG 3038
QY 4386 actcgatcatattgcatgcaaaagaaaaatcccatcctcctcgaagcttcta 4445
Db 3039 ACTCGATGCTATCATTCATGATGAAGAAAAAATCCACATCTCTGCTCAAGACGCTTCTA 3098
QY 4446 ccagctcaccagctcctggactccgctcagcctattgcagagagctgcactgctcac 4505
Db 3099 CCAGCTCACCAAGCTCCTGGATCCGCTGTCAGCTTATTCGAGAGAGCTGCATCAGTTTAC 3158
QY 4506 ttttgactgtcaataagtcacacatggtgagcgtgactttccggaaatgatggcaga 4565
|||||
```

```
Db 3159 TTTTGACCTGCTAATCAAGTCACACATGGTGAGCTGGACTTTCCGGAAATGATGGCAGA 3218
QY 4566 gatcatctctgtgcgaagtgcccaagatccttcttggaagtcgaagcccatctatttcca 4625
Db 3219 GATCATCTCTGTGCAAGTGCCCAAGATCCTTTCGGGAAAGTCAAGCCCATCTATTTCOA 3278
QY 4626 caccagtggaagcattgaaaccctatttcccccacccagctcatgcccccttccagatg 4685
Db 3279 CACCAGTGAAGCAPTGGAAACCCCTATTTCGCCACCCAGCTCATGCCCCCTTTTCAGATG 3338
QY 4686 tcttctgcctgttataactctgcactactcctctgcagctgccttggggaatttccctcat 4745
Db 3339 TCTTCTGCTGTATAACTCTGCACACTCTCTCTGCAGTGGCTTGTTTAAATTTCTCTCAT 3398
QY 4746 tgatgcagctgtctgaacatgttctcctgaattctctatttctgtgggctttttttctc 4805
Db 3399 TGATGTACAGTCTGTCATG-----GAATCTATTGCTGGGCTTTTTTTTTTCTC 3447
QY 4806 ttctctccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctc 4865
Db 3448 TTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3507
QY 4866 ttgtctcccatctgtggcctctctctgttttgaatgtgtgtgtgtgtgtgtgtgtgtgtgt 4925
Db 3508 TTTGCTTCCCATTTGGGCTCCTATCTCTGTGTTTGAATGGTGTGTGTGTGTGTGTGTGTGT 3567
QY 4926 tgatgcctcctcatgcccagtgctcaagttgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4985
Db 3568 TGATGATCCTCATATGGCCAGTGTCAAGTTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 3627
QY 4986 gccacacaaagtttacttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 5045
Db 3628 GCCACACAAAGCTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTATCTGG 3687
QY 5046 ggaaatcaaaacaaaaa 5062
Db 3688 GGAAATCAAAACAAAAA 3704

RESULT 2
LOCUS HUMARB 3715 bp mRNA PRI 31-OCT-1994
DEFINITION Human androgen receptor mRNA, complete cds.
ACCESSION M3263 N18624
VERSION M3263.1 GI:178893
KEYWORDS androgen receptor; dihydrotestosterone receptor.
SOURCE Human prostate, cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3715)
AUTHORS Chang,C.S., Kokontis,J. and Liao,S.T.
TITLE Structural analysis of complementary DNA and amino acid sequences
of human and rat androgen receptors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (19), 7211-7215 (1988)
MEDLINE 89017168
REFERENCE 2 (sites)
AUTHORS Chang,C.S., Kokontis,J. and Liao,S.T.
TITLE Molecular cloning of human and rat complementary DNA encoding
androgen receptors
JOURNAL Science 240 (4850), 324-326 (1988)
MEDLINE 88178111
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by S.Liao, 01-MAR-1989.
Four proteins of different sizes were found. These proteins could
be explained if alternative start codons at positions 1084, 1252
and 1315 were used.
FEATURES
source Location/Qualifiers
1..3715
/db_xref="taxon:9606"
/map="Xq11.2-q12"
532..3288
gene
```

CDS	/gene="AR" 532..3288 /gene="AR" /codon_start=1 /db_xref="GDB:600-120-556" /product="androgen receptor" protein_id="AAA51775.1" /db_xref="GI:178894"
	translation="MEVQLGLRVYPRPSPKTYRGAFONLFQSVREIVNPGPRHPEA ASAAPGASLLILLOOQQOOQQOQQOQTSPROOQQOQEDSGSPQAHRRGPTGYL VLDEEQOPQSPOSALECHPERGCVPCEGAVAASKGLPOOLPAPEDDSDAAFTLSL LGPTFFGLSSDAKDILKSILSASTMOLLQOQQOENVSSEGSRRAREAGAPTSSKDN YLGTSTISDNAKEIKAVSVSMGLGVBEALBHLSPGEOLRDCMTAPILLGVPPAVRPT PCPAIECKGSLILDAGSKSETDIAEYSPFKGGYTTLKEGESLSCSSAAGSSGTLE LPSTLSLYKSGALDDAAAYQRDYNYFLALAGPPPPLPPPHPARIKLENPLYGSA WAAAAOCRYGDLASLHGAGAAGPGSGPSAAASSWHITLFTABEGQLYPCGGGGGG GGGGGGGGGGGGGGGGGGGEEAEAVAPYGVTRPQGLAGOESEDFTADPVWVGGMVS RYPYSPTCVSEMGPWMDSYSGPYGDMRLIETARDHVLPIDYFPPPKTCLICGDEAS GCHYALTCGSKGVFFRKAAEBGKIILCASRNDCITIDFRKRKNFOSCLRKYRCBAGWT LGARKLKGNLUKEFEAEASTTSPTTEOTKLTVSHIEGECOIPELNYLEALEPG VYCAGDNNQPDSEFAALLSNELGERQLVHVYWKAKLPGRFNLDVDQMAVTOYSW MGMLVFAMGVRFTNVNSRMILYFAPDLFVNEYMRHKSRMYSQCYVMRHLSOFELWQI TPQEFLMKALLLFSII PVDLKLNKFDELRLMNYIKELDRILIACRKKNPTSCSRREY OLTKLIDSVPQIARELHQTFDLLLIKSHVSVDPFEMMAELIISVQVPKILSGVKPIY FHTQ"
BASE COUNT	841 a 1055 c 1001 g 818 t
ORIGIN	1 bp upstream of EcoRI site; chromosome Xq11.2-q12.
Query Match	70.8%; Score 3595.8; DB 9; Length 3715;
Best Local Similarity	98.6%; Pred. No. 0;
Matches 3685; Conservative	0; Mismatches 7; Indels 45; Gaps 4;
Qy	1335 tgcaaggagagaacccctgtgtttcccccaactctcttccaacctcctcgttcctccc 1394
Dd	4 TTCCGGCGGAGAGAACCCTCTGTTTTCGCCCACTCTCTCCTCCACCTCCTCGCTTCCC 63
Qy	1395 caccggatgcggagccagagatcaaaagatgaagcgagtcaggtcttcagtagcca 1454
Dd	64 CACCCGAGTGCGGAG-CAGAGATCAAAGNTCAAAAGCGAGTCAGGCTCTTCAGTAGCCA 122
Qy	1455 aaaaacaaaaacaaaaaacgagcggaataaaagaaagataataaactcagt 1514
Dd	123 AAAAACAAAAACAACAAAAAAGCGCGAATAAAGAAAAAGATATAACTCAGTT 182
Qy	1515 ctatttgacctacttcagtgacactgaatttgaagtgagagatttgttttttc 1574
Dd	183 CTATTATGCACTTCTACGTGGACACTGAATTTGAAGTGAGGATTTGTGTTTTTC 242
Qy	1575 tttaagatctgggcatcttttgaattcaccttcgaattattaagagacagactgtgagc 1634
Dd	243 TTTTAAGATCTGGCATCTTTTGAATCTACCTTCAAGTATTAAAGACAGACTGTGAGC 302
Qy	1635 ctagcaggcgagatctttgtccacgtgtgtctctctctgtcacgagactttgaggtgtca 1694
Dd	303 CTAGCAGGCGAGATCTTGTCTCACCGTGTGTCTTCTTCTGACGAGACTTTGAGGCTGTCA 362
Qy	1695 gagcgcttttgtggtgtgtctcccgcaagtctcctctctgtgagcttcccgagtggg 1754
Dd	363 GAGCGCTTTTTCGTGTGCTCCCGCAAGTTTTCCTTCTCTGAGCTTCCCGCAGGTGG 422
Qy	1755 cagctagctgacgactaccgcatacatcacgctgttgaaactctctgtgagcaagaaa 1814
Dd	423 CAGCTAGCTGCAGCGACTACCGATCATATCAGACGCTGTTGAACCTCTTCTGAGCAAGAGAA 482
Qy	1815 gggagcggggttaaggagtaggtggaagattcaagcagctcaagatgaagtga 1874
Dd	483 GGGAGGCGGGGTAAAGGAAGTAGTGGGAAGATTACGCCAAGCTCAAGGATGGAAGTGCA 542
Qy	1875 gttagggctgggaagggtctacctcggccgctccagacacctcacgagagctttcca 1934
Dd	543 GTTAGGCTGGGAAGGCTTACCTTCGGCCCGCTCCACACACTACCGAGGAGCTTTCCA 602

Db	1659	TCCCCACGCTCGCATCAAGCTGGAGAACCGCTGACGTACGGCAGCGCTGGCGGGCTGC	1718
Qy	3075	ggcgccagtgccgctatgggacactggcgagcctgcatggcgcggtgagcgagacc	3134
Db	1719	GGCGGCGCAGTGGCGGTATGGGACACTGCGGAGCCTGATGSCGCGGCTGCAGCGGAC	1778
Qy	3135	cggttcgggtcaacctcagccgagccgctccatctcctctggaacactctctcaagccga	3194
Db	1779	CGGTTCGGGTCAACCTCAAGCGCGCTTCCATCTCTGGCACACTCTCTTCACAGCCGA	1838
Qy	3195	agaagccagctgtatggacgctggtggtggtgggggtggt-----ggcgcgcg	3245
Db	1839	AGAAAGCCAGTTGTATGGACCGTGTGTGTGTGGGGTGGTGTGCGCGCGCGGCGG	1898
Qy	3246	cg	3305
Db	1899	CGGCG	1958
Qy	3306	agccccctacgctacactcgccccctcagggcgctggcgggccagaaagcaacttcac	3365
Db	1959	AGCCCCCTACGGCTACACTCGGCCCTCAGGGGCTGCGGGCCAGGAAAGCAAGCTTCAC	2018
Qy	3366	cgcaactgtagtgtgtgtaacctggcggaatggtgagcagagtgcctatccagctcccac	3425
Db	2019	CGCACCTGTATGTGTGTACCTTGGCGCATGTGTGAGCAGAGTGGCCTATGCCAGTCCAC	2078
Qy	3426	ttgtgtcaaaagcgaatggccccctggatgtagctactccggaccttacggggacat	3485
Db	2079	TTGTGTCAAAGCGAATGGCCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACAT	2138
Qy	3486	gcgtttggagactgccagggaacatgtttggccattgactatttacttccaccccagaa	3545
Db	2139	GCgtTTGGAGACTGCCAGGACCATGTTTGGCCATTGACTATTACTTTCCACCCTCAGAA	2198
Qy	3546	gaactcctgaatctfgagatgaagctctgggtgtcaactatggagctctcaacttgg	3605
Db	2199	GACCTGCTGATCTGTGGAGATGAAGCTTCTGGGTCTCACTATGGAGCTCTCACATGTG	2258
Qy	3606	aagctgaaggtctcttcaaaagagcgctgaagggaacagagaaglacctgtgcgcag	3665
Db	2259	AAGCTGCAAGGTCTTCTTCAAAGAGCGCTGAAGGGAAACAGAAGTACCTGTGCGCAG	2318
Qy	3666	caaaatgattgaactattgataaattccgaaggaaaaattgtccatcttgtctgtcttctg	3725
Db	2319	CAGAAATGATTGCATCTATTGATAAAATCCGAAGGAAAAATTTGCCATCTTGTCTGTCTTCG	2378
Qy	3726	gaaatgttagagcaggatgactctggagcccggaagctgaagaaacttgttaact	3785
Db	2379	GAATGTTATGAAGCAGGATGACTCTGGAGCCCGAAGCTGAAGAAACTTGTGAATCT	2438
Qy	3786	gaaactacaggagggaaggaggtctcagcaccaccgccccactgagagacaacca	3845
Db	2439	GAAACTACAGGAGGAGGAGGCTTCCAGCACCAAGCCCCACTGAGGAGACAACCCA	2498
Qy	3846	gaagctgaagtgtaacacattgaaggtatgaattgtcgagcctcttctggaatgctct	3905
Db	2499	GAGCTGACAGTGTACACATTGAAGGCTATCAATGTGAGCCCCATCTTTCTGAATGTCTCT	2558
Qy	3906	ggaagccattgagccagggtgtagtgtgtgtgtagcaagcaacaacccagccccactctt	3965
Db	2559	GGAAAGCAATTAGCCAGGTGTAGTGTGTCTGGACACGACAAACACAGCCCGACTCTT	2618
Qy	3966	tgcagctgtcctctagcctcaatgaactggagagagacagctgtcacactgtgtcaaa	4025
Db	2619	TGACAGCTTGTCTCTPAGGCTCAATGAATCTGGGAGAGAGAGCTTGTACACTGTGTCAA	2678
Qy	4026	gtgggccaagccttgcctggcttccgaacttacaagtgagcaccagatgctgtcat	4085
Db	2679	GTGGGCCAAGGCTTGCCTTGGCTTCGGCACTTACAGTGGACACCAGATGCTGTTCAT	2738
Qy	4086	tcagtactcctgtaggggctcatggttttgccatgggtggcgatccttccaccaatgt	4145

Db	2739	TCAGTACTCCTGGATGGGGCTCATGTTGTTTGCATATGGCTGGCGATCTCTTCCACCAATGT	2759
Qy	4146	caactccaggatgcttacttcgccctgatctggttttcaatgagtacccagatgcacaa	4205
Db	2799	CAACTCCAGGATGCTTACTTCCGCCCTGATCTGGTTTCAATGAGTACCGGATGCACAA	2858
Qy	4206	gtcccgatgtacagccagtggtgcgaatgaggcaacctctctcaagagtttgatggct	4265
Db	2859	GTCCCGATGTTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTGGATGGCT	2918
Qy	4266	ccaatcaccgccaggaattcgtgcatgaagcaactgactcttcagcattattcc	4325
Db	2919	CCAATCACCCTCCAGGAATCTGTGCATGAAGACACTACTCTTCACGATTATTCC	2978
Qy	4326	agtgatgggtgaaatacaaaattcttggatgaactcgaaTgaactacatacaagga	4385
Db	2979	ACTGGATGGGTGAAAAATCAAAAATCTTTGATGAACCTCGAATGAATACATCAAGGA	3038
Qy	4386	actgatgttatcatgtcatgcaaaagaaaaatccacatctctctcctaagaagcttcta	4445
Db	3039	ACTCGATCGTATCATTTGATGCAATGCAAAAAAGAAAAATCCACATCTCTCAAGACGCTCTA	3098
Qy	4446	ccagctcaccaagctccgtgactccgtgcagcctatttgcgagagagctgcatacagttcac	4505
Db	3099	CCAGCTCACCAGAAGCTCCTGGACTCCGTGCAGCCTATTGGAGAGAGACGTGCATCATGTTCCAC	3158
Qy	4506	tttgacctgtaatacgaatcacatggtgagcgtggactttccggaaatgatggcaga	4565
Db	3159	TTTTGACCTCGTAATCAAGTTCACACATGTTGAGCGTGGAGTTTCCGGAAATGATGGCAGA	3218
Qy	4566	gatcatctctgtcgaagtgcgaagatccttctggaagaaagtcagaccatctattcca	4625
Db	3219	GATCATCTCTGCAAGTGCACCAAGATCCTTTCTGGAAAGTCAAGCCCATCTATTTCCTCA	3278
Qy	4626	caccagtgaaagcatgtgaaacccctatttcccccacccagctcatgccccctttcagatg	4685
Db	3279	CACCCAGTGAAGCATTTGAAACCCCTATTTCGCCACCCAGCTCATGCCCTTTTCAGATG	3338
Qy	4686	tctctgcctgtttataactctgaactactcctctgcagtgcccttggggaatttcccttat	4745
Db	3339	TCCTCTGCCCTGTTATAACTCTGCACCTACCTCTCTCGAGTGGCTTGTAAATTCCTCTAT	3398
Qy	4746	tgatgacagctgtgcatgaacatgttccctgaattctatttctgtgggcttttttttctc	4805
Db	3399	TGATGTACAGTCTGTATG-----GAATCTATTTCGTGGCTTTTTTTTTTCTC	3447
Qy	4806	ttctctccctttcttttcttctccctccctatctaaacctcccatggaaccttcagac	4865
Db	3448	TTTCTCTCCCTTTCTTTCTCTTCCTCCCTCATTAACCCCTCCCATGGACCTTCAGAC	3507
Qy	4866	tttgcttcccattgtggctcctatctgttttgaaatgggtgtgtatgacctttaaattctg	4925
Db	3508	TTTGTCTTCCCATTTGGCTCCTATCTGTGTTTGAATGGTGTGTATGCCCTTTAAATCTG	3567
Qy	4926	tgatgactctcatatggcccaagtgtcaagtdtggcttgttttacagcactactctgtgcc	4985
Db	3568	TGATGATCTCTCATATGGCCAGTGTCAAGTTGTGCTTGTATGCCCTTTAAATCTG	3627
Qy	4986	gcacacaaacgtttacttatcttatgtccacgggaagtttagagagctaaattatctgg	5045
Db	3628	GCCACACAAACGTTTACTTATCTTATGCCCAGGGGAGTTTAGAGAGCTTAAGATTATCTGG	3687
Qy	5046	ggaatcaaacacaaaaa 5062	
Db	3688	GGAAATCAAAACAAAAA 3704	
RESULT	3		
LOCUS	109479		
DEFINITION	Sequence 1 from Patent WO 8909223.	PAT	02-DEC-1994
ACCESSION	109479		
VERSION	109479.1	GI:587808	

RESULT	3		
I09479			
LOCUS	I09479	3721 bp	PAT
DEFINITION	Sequence 1 from Patent WO 8909223.		
ACCESSION	I09479		
VERSION	I09479.1	GI:587808	

[illegible][illegible]

RESULT		4	PAT	02-DEC-1984
I09510				
LOCUS	T09510	3569 bp		
DEFINITION	Sequence 1 from Patent WO 8909791.			
ACCESSION	I09510			
VERSION	T09510.1 GI:587782			
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown..			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 3569)			
JOURNAL	French,F.S., Wilson,E.M., Joseph,D.R. and Lubahn,D.B. DNA ENCODING ANDROGEN RECEPTOR PROTEIN Patent: WO 8909791-A 1 19-OCT-1989;			
FEATURES	Location/Qualifiers l...3569 /organism="unknown" 796 a 1009 c 974 g 790 t			
BASE COUNT				
ORIGIN				

Query Match 69.8%; Score 3547; DB 6; Length 3569;									
Best Local Similarity 99.7%; Pred. No. 0;									
Matches 3569; Conservative 0; Mismatches 0; Indels 12; Gaps 1;									
QY	1502	taataactcagttcttatttcacactaacttcagtgagacactgaaatttgaaaggtggagga	1561						
Db	1	TANTACTCAGTCTTATTGTGACCTTACTAGTGACACTGAATTTTGAAGGTGGAGGA	60						
QY	1562	ttttgtttttttttttaagatctgggcattcttttgaatctacccttcaagattattaagag	1621						
Db	61	TTTTGTGTTTTTCTTTTAAGATCTGGGCATCTTTGAATCTACCCCTTCAAGTATTAGAG	120						
QY	1622	acagactgtgagccttagcaggcagatcttgcacccgtgtgtcttcttcgcacgagac	1681						
Db	121	ACAGACTGTGAGCCTTAGCAGGCGCATCTGTGCCACCGTGTCTCTCTCTGCACGAGAC	180						
QY	1682	tttgaggtcgtcagagcgttttttgcgtgtgtgtctccgcgaagttctctctctcgtgagct	1741						
Db	181	TTTGTAGGCGTGCAGAGCGCTTTTTCGGTGGTGTCTCCGCAAGTTTCTTCTCTGGAGCT	240						
QY	1742	tcccgaggtggcgagctagctgcgcgactacogcatcatcagcgcgtgttgaaactctt	1801						
Db	241	TCCCGCAGGTGGGCAGCTAGCTGTCAGCGACTACCGCATCATACAGCGCTGTGAACCTCT	300						
QY	1802	ctgagcaagagaaggggagcggggtcaaggaagtagtggaagatcagccaagctcaa	1861						
Db	301	CTGCAGCAAGAGAGGGAGCGGGGTAGGGAAGTAGGTGGGAATTTTCAGCAAGCTCAA	360						
QY	1862	ggaatgaagtgtagttagggctggaaagggctacacctgcgcgcgcgtccaaagacctaac	1921						
Db	361	GGATGGAGTGCAGTTAGGGCTGGGAAGGGTCTACCCCTCGCGCGCTTCCAAAGACCTACC	420						
QY	1922	gagggacttccagaatctgttccagaacgtgcgcgaagtgcacgaacccggggcccca	1981						
Db	421	GAGGAGCTTTCAGAAATCTGTGTCCAGAGCGTGCGCGAAGTGATCCGAAGCCGGGGCCCCA	480						
QY	1982	ggcaccacagagccgcgagcgcagcacctccggcgccagtttgtctgctgctgcagcgc	2041						
Db	481	GGCACCCAGAGGCCCGCAGCGCAGCACCTCCCGCGCCAGTTTGTCTGCTGCT-----	532						
QY	2042	agcacagcacagcacagcacgacgacgacgacgacgacgacgacgacgacgacgacgacg	2101						
Db	533	----CCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	588						
QY	2102	agcagaagagactagccccagggcagcagcagcagcagcagcagcagcagcagcagcagc	2161						
Db	589	AGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	648						
QY	2162	aagccatctgtagggcccaacaggtacctgtctctgtagtgggaacagcaaccttcac	2221						
Db	649	AAGCCCATCTGTAGAGGGCCCCACGGCTACCTGTCTCTGTGATGAGGAACAGCAACCTTCAC	708						
QY	2222	agccgagtcggccctggagtgcacacccgagagagttgcgtccagagccttgagccg	2281						
Db	709	AGCCGCAGTCGGCCCTTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCGAGAGCCTTGAGCCG	768						
QY	2282	ccgtggccgcagcaagggctgcgcgagcagctgcagcacctccggacgaggtgact	2341						
Db	769	CCGTGGCCGCCGAGCAAGGGGCTGCCGAGCAGCTGTCAGCACCTCCGGACGAGGATGACT	828						
QY	2342	cagctgccccatccagcttgtcctgctgggccccactttcccgacttaagcagctgct	2401						
Db	829	CAGCTGCCCCATCCAGTTGTCTCTGTGGGCCCCACTTTCCCGCGCTTTAAGCAGCTGCT	888						
QY	2402	ccgtgaccttaagacatccttgagcgaagccagaccatgcaactccttcagcaacagc	2461						
Db	889	CCGTGACCTTTAAGACATCTCTGAGCGAGGCCAGCACCTGCAACTCTTCAGCAACAGC	948						
QY	2462	agcagggaagcagtatccgaagcagcagcagcgggagagcggagggagggcctcggggctc	2521						
Db	949	AGCAGGAAGCAGTATCCGAGGCGAGCAGCGGGAGAGCGGGAGGAGGCTTCGGGGGCTC	1008						

QY	2522	ccacttctccaaaggacaatttacttagggggcacttcgaccatttctgacaaacgccaagg	2581
Db	1009	CCACTTCTCTCCAAGGACAAATTACTTAGGGGGCACTTCGACCATTCTGACAAGCCAAAG	1068
QY	2582	agttgtgtaaggcagtgctcgctgtccatgggacctgggtggagggcgtttggagcaactga	2641
Db	1069	AGTTGTGTAAAGCAGCTGTCGGGTGTCATATGGCTTGGTGTGAGGCGTTTGGAGCATCTGA	1128
QY	2642	gtccaggggaacagctctcgggggattgcattgacgccccacttttggagttccaccgcg	2701
Db	1129	GTCCAGGGGAACAGCTTCGGGGGATTGCAATGATACGCCCACTTTTGGGAGTTCCACCCG	1188
QY	2702	ctgtcggtcccaactcttctgtgccccattggccgaatgcgaaagtgcgaagttctctgtacagaca	2761
Db	1189	CTGTGGTGCCCACTTCCTTTGTGCCCCATTTGGCCGAATGCAAGAGTTCTCTGTGTAGACGACA	1248
QY	2762	gcgcagggaagcagctgaagatactgctgagtattcccccttccaagggaggttaccacca	2821
Db	1249	GGCGAGGCAAGAGCACATGAAGATACTGCTGAGTATTCCTTTTCAAGGGAGGTTTACACCA	1308
QY	2822	aaaggctagaaggcgaagagcctaggctgtcttggcagcgctgcagcaggaggagctccggga	2881
Db	1309	AAGGGCTATGAAGGCGAGAGCCTTAGGTGCTCTGGCAGCGCTGCAGCAGGAGTCCGGGA	1368
QY	2882	cacttgaactgccgtctaccctgtctctctacaagtccgagacactggacagagcactg	2941
Db	1369	CACTTGACTGCCCTCTACCCCTGTCTCTTACAGTTCGGAGCACTTGCACGAGGACGCTG	1428
QY	2942	ggtaccagagtcgcgactacatacaactttccactggctcttgccggacgcgcgccctc	3001
Db	1429	CGTACGAGATFCGCGACTACTACAATTTCCACTGGCTCTGGCGGACCGCCGCCCTC	1488
QY	3002	gcgcgctccccattccccacgctcgcatcaaatggatggagaaacccgctggactacggcagcg	3061
Db	1489	CGCGCCTTCCCATTCGCCACGCTTCGCATCAAGCTGGAAACCCGCTGCATTCGGCAGCG	1548
QY	3062	cctgggcggtgcgggcgcagtcgcgcta tggggacctggcgagcctgcattgggcgcg	3121
Db	1549	CTGCGGGGCTGCGGCGGCGCAGTGCCTATATGGGACCTTGGCAGCCTTCATGGCGCGG	1608
QY	3122	gtgcagcggaaccggttcttggllcaccctcagccgcgcttccctatcctgtgcacactc	3181
Db	1609	GTGCAGCGGAGCCCGTCTGGGTCAACCTCAGCGCGCGCTTCCTCATCTGSCACACTC	1668
QY	3182	tcttcacagccgaaggaagccagttgtat tggaccgctggtgtgggtgggggtgggtggcg	3241
Db	1669	TCATTACAGCCGAAGAAGCCAGTTGTATATGGAACCGTGTGTGTGGTGGGGGTGTGGCG	1728
QY	3242	gcgcgcggcg	3301
Db	1729	CGGAG	1788
QY	3302	ctgtagccccctacgctacactcggccccctcaggggctggcgggcgacgaagaaagcact	3361
Db	1789	CTGTAGCCCCCTTACGCGTACTTCGSCCCCCCTCAGGGGCTTGGGGGCCAGGAAGCGACT	1848
QY	3362	tcaccgcacctgatgtgtgtaccctggcggtat tggtagcagagtgcacctatccacgtc	3421
Db	1849	TCACCGCACTGATGTGTGTGTATACCTTGGCGCATGTTGTGACAGAGTGCCTATATCCAGTC	1908
QY	3422	ccacttgttcaaaagcaaaatggggccccctgagtagtagactactccgacacctacgggg	3481
Db	1909	CCACTTGTCTCAAAAGCGAATATGGGCCCTTGGATGGATAGCTACTCCGAGACCTTACGGGG	1968
QY	3482	acatgcgtttgagactgccagggaacctatgtttggccattgactattactttccacccc	3541
Db	1969	ACATGCGTTTGGAGACTGCCAGGACCAATGTTTGGCCATTGACTATTATTTTCCACCCC	2028
QY	3542	agaagacctgcctgactcttgagataagaactctcgggtgtcaactatgagctctcaat	3601
Db	2029	AGAAGACCTTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGAGGCTCTCAAT	2088
QY	3602	gtggaagctgcagggtcttcttcaaaagacgcgctgaagsgaagaaacgaagatcctgtgcg	3661

GVYCAHDNINPDFAALLSSINELGERQLVHVVKAKLPGRNLHWDQMAVIOYS
WMLMVFAMGRNSTNNSRMLYFAPDLVFNEXRBMKSRMYSCQVRNHL SOEFGWLO
ITPOEFLCMKALLIFSITPDGLKNOKFEDELNNYJKELDRIITACKRKPSTCSRRE
YOLTKLDSVOPiARELHOFTFDLLIKSHMVSVDFPEMMAEIIISQVPKILSGKRVPi
YFHQO"

BASE COUNT 796 a 1009 c 974 g 790 t
ORIGIN

Query Match 69.8%; Score 3547; DB 9; Length 3569;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3569; Conservative 0; Mismatches 0; Indels 12; Gaps 1;
QY 1502 taataactcagttcttattgcaccactcttcagtgagacactgaatttggaaaggtggagga 1561
Db 1 TAATAACTCAGTTCATTATTGACCTACTTCAGTGGACACTGAATTTGAAGAGTGGAGGA 60
QY 1562 ttttggtttttttttaagatctgggcatcttttgaatctacoccttcaagattaaag 1621
Db 61 TTTTGTGTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCGCTTCAAGTATTAAAG 120
QY 1622 acagactgtgagctagcagggcagatctgtccacgtgtgtcttcttcttcttctggagct 1741
Db 121 ACAGACTGTGACCTTAGCAGGCGAGATCTTGCCACCGGTGTCTTCTTCTGSCACGAGAC 180
QY 1682 tttgaggtgtcagagcgtcttttgcgtgtgttgcctccgcgaagtcttcttcttctggagct 1741
Db 181 TTTGAGGCTGTGACAGCGCTTTTTCGCTGGTGTGTCGCCGAAGTTTCTTCTCTGGAGCT 240
QY 1742 tccccaggtggcagctagctgcagcagactaccgcataccacagcctgttgaactctt 1801
Db 241 TCCCCAGGTGGCAGCTTAGCTGACGACTACCGCATCATCACAGCGCTGTTGAACCTTT 300
QY 1802 ctgagcaagagaaggggagcgggtaaggaagtaggtggaagattcagccaagctcaa 1861
Db 301 CTGACCAAGAGAAGGGAGCGGGTAAGGAAGTAGGTGGAAGATTACGCCAAGCTCAA 360
QY 1862 ggaatgaagtcagttaggtggaaggggtctaccttcggcccgctccaaagacctacc 1921
Db 361 GGATGGAAGTCAGTTAGGGTGGGAAGGGTCTACCGTCCGCCCGCTCCAAAGACCTACC 420
QY 1922 gaggagcttccagaatctgtccagagcgtgcgcaagtgatccagaacccgggcccga 1981
Db 421 GAGGAGCTTTCAGANATCTGTTCAGAGCGTGC CGAAGTATCCAGAACCCGGGGCCCCA 480
QY 1982 ggcacccagggccgcagagcagcaactccggcgccagtttgctgctgtgcagcagc 2041
Db 481 GGCACCCAGAGCGCGAGCGACGACCTCCCGCGCCAGTTTGCCTGCT 532
QY 2042 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2101
Db 533 -----GCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 588
QY 2102 agcagcaagagactagccccagggcagcagcagcagcagcagcagcagcagcagcagc 2161
Db 589 AGCAGCAAGAGACTAGCCCCAGGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 648
QY 2162 aagcccatcgtagagggcccaagggctacctggtcctggatgaggaacagcaaccttcac 2221
Db 649 AAGCCCATCTGATAGAGGCCACAGGCTACCTGTGCTTGGATGAGGAACAGCAACCTTCAC 708
QY 2222 agccagcagtcggccctagagtgcacccccagagagaggttgcgtccacagagcctggagccg 2281
Db 709 AGCCGCACTG66CCCTGGAGTGGCCACCCCGAGAGAGGTTCCGTCCAGAGGCTTGGAGCGG 768
QY 2282 ccgtggccgcagcaaggggtgcgcagcagctgccagcactcccgagcagagtagact 2341
Db 769 CCGTGGCCGCCAGCAAGGGGCTGCCGAGCAGCTGCCAGCACCCTCCGGAGGAGTATGACT 828
QY 2342 cagctgcccatccaaagtgttccctgtggggccccactttccccggcttaagcagctgct 2401
Db 829 CAGCTGCCCATCCACAGCTTGTCCGTGTGGGGCCCCACTTTCCCGCGCTTAAGCAGCTGCT 888

QY 2402 ccgctgaccttaagacacatcctgagcgagggccagcaccatgcgaactccttcagcaaacagc 2461
Db 889 CCGGTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGC 948
QY 2462 agcaggaagcagLatcccaagagcagcagcagcgagggagagcgagcctcgggggctc 2521
Db 949 AGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCTCTGGGGGCTC 1008
QY 2522 ccacttctccaaggaacaattacttaggggggaacttcgaccatttctgacaaagcgaag 2581
Db 1009 CCACCTTCTCAAGGACAATTAATTAGGGGGCACTTCGACCATTTCAGCAACGCCAAG 1068
QY 2582 agttgttaagcagctgcggttccatagggcctgggtgagggcgttggagacatctga 2641
Db 1069 AGTTGTGTAAGGCAGTGTGGTGTCCATGGCGTGGGTGTGGAGCGCTTGGAGCATCTGA 1128
QY 2642 gtccaggggaacagcttcgggggattgcatgtacgcccccaacttttgggaggttccaccg 2701
Db 1129 GTCCAGGGGAACAGCTTCGGGGGATTTGCATGTACGCCCACTTTTGGAGGTTCACCCG 1188
QY 2702 ctgtgcgtccactcctgtgtcccaattggcccaattggcgaatgcgaagttctctgtagacga 2761
Db 1189 CTGTGCTGCCACTCTCTTGTGCCCAATTGGCGGAATGCAAGGTTCCTCTGCTAGACGACA 1248
QY 2762 ggcagggcaagcagcactgaagatactgctgagttattcccttccaaagggaggttaacca 2821
Db 1249 GCGCAGGCAAGACACTGAAGATACGTGCTGAGTATTTCCCTTTCAAGGGAGGTTACACCA 1308
QY 2822 aagggctagaagcgagagccttaggctgctgtgagcgtcagcagcgagctcccgga 2881
Db 1309 AAGGGCTAGAAGCGAGAGCCTTAGGCTGTCTTGGCAGCGCTGCAGCAGGAGCTCCGGGA 1368
QY 2882 cacttgaactgcgtgtacacctgtctctctataagtcggagcactgcagcagcgagctg 2941
Db 1369 CACTTGAATGCGCGTCTACCCCTGTCTCTACAGTCCCGAGCACTGGACGAGGAGCTG 1428
QY 2942 cgtaccagagtcgcgactactacaactttccactggctgctgcccagcagcgcccccctc 3001
Db 1429 CGTACCAGAGTCGCGACTACTACAACTTTCACACTGGCTGTGCGCGAGCGCGCCGCCCTC 1488
QY 3002 cgcgcctcccaatcccaacgctcgcataagctggagaacccgcctggactacggcagcg 3061
Db 1489 CGCGCGCTCCCATCCACGCTCGCATCAAGTGGAGAAACCCGCTGACTACGCGACGC 1548
QY 3062 cctggcgcgctgcgcgcagtcagctccgctatggagacactggcagcagcctgcattggcgcg 3121
Db 1549 CCGTGGCGGCTCGGGGGCGCAGTCCCGTATGGGAGCACTGGCGAGCCTGATGGCGCGG 1608
QY 3122 gtgcagcggggacccggttctgtgggtcacacctcagcccgcttccctcatcctggcacactc 3181
Db 1609 GTGCAGCGGGACCCGTTCTGGGTCAACCTCAGCGCGCGCTTCCTCATCTTGGCACACTC 1668
QY 3182 tcttcacagcgaagaagccagttgtatggaccggtggtggtgggtgggtgggtgggtg 3241
Db 1669 TCTTACAGCCGGAAGAGCCAGTTGTATGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1728
QY 3242 ggcggcg 3301
Db 1729 GCG 1788
QY 3302 ctgtagcccccttaeggtctacactgcgccccctcaggggctggcgggcgaggaagcagct 3361
Db 1789 CTGTAGCCCCCTACGGCTACACTTGGCCCCCTCAGGGGCTGCGCGCCAGGAAGCCACT 1848
QY 3362 tcacgcacctgt 3421
Db 1849 TCACGCGCACTGTATGT 1908
QY 3422 ccaactgtgtcaaaagcgaaatggcccttgatggatagctactccgaccttaccggg 3481
Db 1909 CCACCTGTGTCAAAAGCGAAATGGGCCCTCGATGGATAGCTACTCGGACCTTACGGGG 1968
QY 3482 acatgctgttgagactgccagggaccatggttttggccattgactattactttccacccc 3541

|||||
 Db 1969 ACATCGCTTTGGAGACTGCCAGGACCAATGTTTGGCCATTGACTATTACTTTCACCCC 2028
 Qy 3542 agaagacctgctgatctgtggagatgaagctctctgggtgctcaetataggagctctcaat 3601
 Db 2029 AGAAGACCTGCGTGTATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCAAT 2088
 Qy 3602 gtggagctgcaaggtctcttcaaaagagccgctgaagggaacagaaglacctgagcg 3661
 Db 2089 GTGGAAGCTGCAAGGCTCTTCAAAAAGAGCGCTGAAGGGAACAGAAAGTACCTGTGCG 2148
 Qy 3662 ccagcagaaatgatgctactattgataatccgaaggaataattgtccatcttgtgtgc 3721
 Db 2149 CCAGCAGAAATGATTGCACTATTGATAAATCCGAAGGAATAATGTCCATCTTGTGCTC 2208
 Qy 3722 ttccgaaatgttatgaagcaggatgaactctgggagcccggaagctgaagaaacttgta 3781
 Db 2209 TTCGGAATGTTATGAAGCAGGGATGACTCTGGGAGCCGGAAGTGAAGAACTTGGTA 2268
 Qy 3782 atctgaaactacaggagaagagaggtccagaccaccagcccaactgagagacaa 3841
 Db 2269 ATCTGAAACTACAGAGGAAGAGAGGCTTCAGCACCACCAGCCCACTGAGGAGACNA 2328
 Qy 3842 ccagaagctgacagtgctacacatgaaggctatgaatgtcagccctcttcttgaatg 3901
 Db 2329 CCCAGAAGCTGACAGTGTACACATGGAAGGCTATGAATGTACGCCCATCTTCTGGAATG 2388
 Qy 3902 tctggaagccattagccaggtgtagtgtgtctggacacacaaacacagcccgact 3961
 Db 2389 TCTGGAAGCCATTGAGCCAGGTGTAGTGTCTGGACACACAAACACCCAGCCGACT 2448
 Qy 3962 ctttgacacctgtctctagcccaatgaactggagagagacagcttgtacacgtgg 4021
 Db 2449 CTTTGCAGCCTTGTCTCTAGCCTCAATGAACGTGGAGAGAGACGCTTGTACACGTGG 2508
 Qy 4022 tcaagtgggccaagccttgctgcttcgcgaacttacacgtggaccagatggctg 4081
 Db 2509 TCAAGTGGGCCAAGGCTTGCCTGGCTTCGCCAACTTACACGTGGAGCACCAGATGGCTG 2568
 Qy 4082 tcattcagttactctggaagggtctatggtgtgtgtccatgggtgggagatccttcacca 4141
 Db 2569 TCATTTCAGTACTCCTGGATGGGCTCATGGTGTGTGCCATGGGCTGGCGATCCTTCACCA 2628
 Qy 4142 atgtcaactcagagctctactctgcacctgactctgttttcaatgtagtaccgcatgc 4201
 Db 2629 ATGTCAACTCCAGGATGCTCTACTTCGCCCTGATCTGGTTTCAATGATGATACCGCATGC 2688
 Qy 4202 acaagtcocggatgtacagccagtggtccgaatgaggcacctctctcaagagtttggat 4261
 Db 2689 ACAAGTCCCGGATGTACAGCCAGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGAT 2748
 Qy 4262 ggcctccaaatcaaccccccaggaaattcctgtgcatgaaagacactgtactcttcagcatta 4321
 Db 2749 GGCTCCAAATCACCCCCAGGAATCTGTGCATGAAGAACACTGTCTACTCTTCAGCATTA 2808
 Qy 4322 ttccagtggtatggctgaaatacaaaaattcttggatgaactcgaatgaactacatca 4381
 Db 2809 TTCCAGTGGATGGCTGAAAAATCAAAAATCTTGTGATGACTTCCGAATGACTACATCA 2868
 Qy 4382 aggaactcgatcgatcattatgcatgcaaaagaaaaataccacatcctgctcaagacgt 4441
 Db 2869 AGGAACTCGATCGTATCATTTGCATGCAAAAAGAAAAATCCACATCCTGCTCAAGACGCT 2928
 Qy 4442 tctaceagctcacaagcctctgactccgtgcagctattgtcgagagagctgacatcagt 4501
 Db 2929 TCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGGCTATTGGCAGAGAGCTGCATCAGT 2988
 Qy 4502 tcactttgacctgctaataatcaagtccacatggtgagcgtgggactttccggaatgatgg 4561
 Db 2989 TCACCTTTGACCTGCTAATCAAGTCACACATGCTGAGCGTGGACTTCCGGAATGATGG 3048
 Qy 4562 cagagatcattctgtgcaagtgcccaagatcctttctggtggaagtcgaagcccatctatt 4621
 |||||

Db 3049 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGAAAGTCAAGCCCATCTATT 3108
 Qy 4622 tcaacacccagtgaaacattgaaacccctatttccccccccagctcatgccccctttca 4681
 Db 3109 TCCACACCCAGTGAAGCATTTGGAACCCCTATTTCGCCACCCAGCTCATGCCCTTTCA 3168
 Qy 4682 gatgtctctgcctgttataaactctgactactcctctcagtgccctggggaattcct 4741
 Db 3169 GATGTCTTCGCTGCTTATAAATCTGCACCTACTCTCTGCAGTGCCTGGGAATTCCT 3228
 Qy 4742 ctattgatgtacagctgtctgaacatgttccgtgaattctatttctgtgggttttttt 4801
 Db 3229 CTATTGATGATACAGTCTGTATGAACATGTTCTGAATCTATTTCCTGGGCTTTTTTT 3288
 Qy 4802 tctcttc 4861
 Db 3289 TCTCTTCT 3348
 Qy 4862 agactttgcttcccatgtgtgctcctctctctctctctctctctctctctctctctct 4921
 Db 3349 AGACTTTGCTTCCCATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3408
 Qy 4922 tctgtgatgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4981
 Db 3409 TCTGTGATGATCCTCTCATATGCGCCAGTGTCAAGTGTCTGTTGAATGGTGTGTATG 3468
 Qy 4982 gccagccacacaaacgtttacttatctatgcccagggaaagttagagagactaagatt 5041
 Db 3469 GCCAGCCACACAAAGTTTACTTATCTTATGCCCAGGGAAGTTTAGAGAGCTAAGAT 3528
 Qy 5042 ctgggggaattcaaaacaaaaaacaaagcaaaacaaaaaa 5082
 Db 3529 CTGGGGAATCAAAACAAAAAACAAGCAACAAAAA 3569

RESULT 6
 HSDJ80804/c
 LOCUS
 DEFINITION
 Human DNA sequence from clone RP4-80804 on chromosome Xq11.1-12. Contains the 5' end of the AR gene for androgen receptor (dihydrotestosterone receptor), a PABPN1 (poly(A)-binding protein, nuclear 1) (PABP2) pseudogene, ESTs STSS, GSSs and two putative CpG islands, complete sequence.
 AL049564
 AL049564.10 GI:4902757
 HTG; androgen receptor; AR; CpG island; PABP2; PABPN1.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 139033)
 Chapman, J.
 Direct Submission
 Submitted (03-APR-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On May 27, 1999 this sequence version replaced gi:4757056.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WormPEP; Information on the WormPEP database can be found at

repeat_region /note="LIME1 repeat: matches 4835. .6158 of consensus" 39987. .40530

repeat_region /note="L2 repeat: matches 2130. .2710 of consensus" 40676. .40809

repeat_region /note="L2 repeat: matches 2579. .2710 of consensus" 40945. .41137

misc_feature /note="L2 repeat: matches 1990. .2185 of consensus" complement(41397. .41743)

misc_feature /note="match: GSS: Em:AQ087063" complement(41397. .41704)

misc_feature /note="match: GSS: Em:AQ091315" 41818. .42458

misc_feature /note="match: GSS: Em:AQ357243" 41819. .42063

repeat_region /note="match: GSS: Em:AQ878802" 42045. .42079

repeat_region /note="L2 repeat: matches 2641. .2675 of consensus" 42132. .42421

Query Match 62.7%; Score 3185.4; DB 9; Length 139033;

Best Local Similarity 97.7%; Pred. No. 0;

Matches 3428; Conservative 0; Mismatches 40; Indels 40; Gaps 19;

Qy 1 gagctctggacaaaattgagcgctatgtgtacatggcaagtggttttttagtggttgtg 60

Db 36813 GAGCTCTGGACAAATTGAGCGCTATGTGTACATGCAAGTGTTTTAGTGTTCGTG 36754

Qy 61 tttaacctgtctgtgggtgatttgcctttgagagctctggatgagaaatcatggttaa 120

Db 36753 TTTACCTTCCTGTGCGGTGATTTTGCCTTTGAGAGTCTGGATGAGAAATGATGTTAA 36694

Qy 121 aggaattccagacagaaagagagagagagagagagagagagagagagagagagag 180

Db 36693 AGGCAATTCAGACAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 36634

Qy 181 ggctgaggttctctagagcaaatggcacaatgccagagagccgagctctatccctatgac 240

Db 36633 GGCTGAGGGTTCTTAGAGCAATGGCACATGCCACAGAGCGCGATCTATCCCTATGACG 36574

Qy 241 gaactctaaagtttcagactaactctctgtgctgtgctgctgctgctgctgctgctgct 300

Db 36573 GAA-TCTAAGGTTTCAG---CAAGTATCTGTGCTGGCTTGGTCA--TGGCTTGTCCCTCAGT 36520

Qy 301 tttagagagactctccactctccactctccactctccactctccactctccactctccactcc 360

Db 36519 TTGTAGGAGACTCTCCACACTCTCCACTCTCCACTCTCCACTCTCCACTCTCCACTCTCC 36460

Qy 361 tggcnagccaggagcnaggtattentatcgtctctctctctctctctctctctctctctngc 416

Db 36459 TGGC-AGCCAGGAGC-AGGTATTCTATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36402

Qy 417 ngttgnttttttagattggnottngnaacaaaatt---tgtatgctggcctccaggaaatc 473

Db 36401 GTTGGTTTTTTAGATTGGGCTTTGGAAACCAAAATTTGGTGTAGTGTGCTGGCTTCCAGGAAATC 36342

Qy 474 tggag-cctggcgctaaaccttggtttagaaagcagagagctattcaggaaagca-ggggt 531

Db 36341 TGGAGCCCTGGCGCTAAACCTTGGTTTAGAAAGCAGAGAGCTATTACAGAAAGAGGGGT 36282

Qy 532 cctccagggttagagctagcct 590

Db 36281 CCTCCAGGGCTAGAGCTAGGCT 36222

Qy 591 ccaagacnacttagcagcgcttagcgcgcgtgagggagagggagagagagagagagagagag 650

Db 36221 CCAAGGCACCTAGGACGGCGTTAGCGCGGTGAGGGGGAGGGGAGAGAGAGAGAGAGAGAG 36162

Qy 651 gggagggaaaagaggttgggaaggaagagagagagagagagagagagagagagagagagag 710

Db 36161 GGGAGGGAAAGAGAGGTGGGAAGCAGAGGAGCGCGCCCGGTGGGGCGGGACCCGACTC 36102

Qy 711 gcannaactgttgattgtctctccactccagcgcgcctccagctccagctccagctccagctcc 770

Db 36101 GCA--AACTGTTGCATTGCTCTCCAGCTCCAGCGCCCTCCAGAGATCCCGGGGAGCC 36044

Qy 771 agcttgctggagagcgggaacggtccggagcaagccagagagcagagagagagagagagag 830

Db 36043 AGCTTGCTTGGGAGAGCGGG-ACGGTCCGGAGCAAGCCAGAGGAGAGGAGGACAGAG 35985

Qy 831 ggaataagga::ccnagctagcgcctccagtgctgtacagnagccgaa-ggacgcaccacg 889

Db 35984 GGAAAAAGGG-CCGAGCTAGCCGCTCCAGTGTGTACAGGAGCCGAAGGAGCAGCCACC 35926

Qy 890 ccagccccagcccggctccagcagacgnaacgcctcttgc-----ngogtctcgaagc 943

Db 35925 CCAGCCCCAGCCGCTCCAGCGACAGCCAAAGCGCTCTTGCGAGCGCGCGCTTCGAAGC 35866

Qy 944 cgcgcgcgagagctgcctctctctctcgtgaagtctttaaagagctctaaagactcgg 1003

Db 35865 CGCGCCCGGAGAGCTGCCCTTCTCTCGGTGAAGTCTTTAAAGCTGCTAAAGACTCGG 35806

Qy 1004 aggaagcaagaaagtgcctggttaggactgacggtgccttctctctctctctctctctccacc 1063

Db 35805 AGGAAGCAAGAAAGTGCCTGTTAGGACTGACGGTGCCTTTGCTCTCTCTCCACC 35746

Qy 1064 cgcctcccccaacctgccttccccctccccctccccctccccctccccctccccctcccc 1123

Db 35745 CGCGCTCCCGCCACCGCTGCCCTTCCCGCTCTCCCGCTCTCTCTCCCGAGCTGCCTCAG 35686

Qy 1124 tcggcccgagcgtgntgncgncgagtttcagagagtaactcccccttttgctgcgagcggg 1243

Db 35685 TCGGCTACTCTCAGCAACCCCGCTACACCGCTCTCCCCACCGCCCCCGCCCCCG 35626

Qy 1244 cagactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1302

Db 35565 CGAG-CTAGCTGCATTTGCAAGAGGCTCTTAGAGCCAGGGAGTCTGGGAGCGGCTT 35507

Qy 1303 cagactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1362

Db 35506 CAGCACTCAGCCAGCGAGCGCGCTGGTGTAGCGTGCACCGGAGAGAACCTCTGTTTTC 35447

Qy 1363 cccactctctccacct 1422

Db 35446 CCCACTCTCTCCACCCT 35387

Qy 1423 agatgaaaggcagtcaggtctcagtagcagcaaaaaacaaacaaacaaacaaacaaacaa 1482

Db 35386 AGATGAAAGGCGAGTCAGGTCTTCAGTAGCCAAAAACAAACAAACAAACAAACAAAG 35327

Qy 1483 ccgaaataaagaaaaagataaactcagttcttatttgacactactctcagtggaact 1542

Db 35326 CCGAAATAAAGAAAAAGATAATAACTCAGTCTCTATTGTCACCTACTCTCAGTGACACT 35267

Qy 1543 gaatttgaagtgaggatatttggtttttttctttaaagatctgggcatcttttgaatct 1602

Db 35266 GAATTTGGAAGTGAGGATTTTGTCTTTTCTTTTAAAGTCTGGGCATCTTTTGAATCT 35207

Qy 1603 accctcaagtataagagacagactgtgacctagcagggcagagatcttgtccaccggt 1662

Db 35206 ACCCTTCAAGTATTAAAGACAGACTGTAGCTAGCAGGGCAGATCTTGTCACCGTGT 35147

Qy 1663 gttctctctcgcagcagactttgagcgtcagagcgtttttgcgtggtgctcccgca 1722

Db 35146 GTCTTCTCTGCAGAGACTTTGAGGCTGTACAGCGCTTTTGTGCTGGTGTCTCCGCA 35087

Qy 1723 agttctctctctgagacttcccgcaggtgggcagctagctgcagcagctaccgcatcat 1782

Db 35086 AGTTTCTCTCTGAGAGTTCGCCGAGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 35027

Qy 1783 cacagcctgtgaactctctcagcaagaagagagagagagagagagagagagagagagag 1842

Db 35026 CACAGCTCTTGAACCTCTTCTGAGCAAGAGAGGGGAGGGGGTAAAGGAAGTAGGTGG 34967

```
QY 1843 aagattcagcaagctcaaggtgaaagtgcagtttaggtggaaggttctacccctcgg 1902
Db 34966 AAGATTCAAGCAAGCTCAAGGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCCTCG 34907
QY 1903 ccgcgcgtccaaagacctaccaggagctttccagaatctgttccagagcgtgcgcgaagt 1962
Db 34906 CCGCGCTCAAGAGCTACCGAGAGAGCTTTCAGAAATCTGTTCAGAGCGTGCAGGAAGTG 34847
QY 1963 atccagaacccggcccccagaccacccagagccgcgcagcagcacctcccgccagct 2022
Db 34846 ATCCAGAACCCGGGCCCCAGGACCCAGAGCCCGGAGCGAGCACCTCCCGGCGCCAGT 34787
QY 2023 ttgctgtctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2082
Db 34786 TTGCTGTCTGCT-----GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 34733
QY 2083 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2142
Db 34732 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 34673
QY 2143 ggtgaggtgttctccccaaagcccatcgtagagggcccccacagctacctggtcctgagt 2202
Db 34672 GGTGAGGTGTGTTCTCCCAAGCCCCATCGTAGAGGCCCCACAGGCTACCTTGGTCTGGAT 34613
QY 2203 gaggaacagcaaccttcacagccgcagctgcgcctgagtgccacccccgagagaggttgc 2262
Db 34612 CAGGAACAGCAAACTTCACAGCCCGCATGTCGGCCCTGGAGTGCCACCCGAGAGAGTTGC 34553
QY 2263 gtccagagcctgagcgcgcgtggtccgcagcaaggggtgcgcgagcagctgcccagca 2322
Db 34552 GTCCAGAGCCTGAGCGCGCGGTGGCCGCCAGCAAGGGGTGTCGCAGCAGCTGCCAGCA 34493
QY 2323 cctccgagcagagtgactcagctgccccatccacgtgtccctgtcgtggtgcccactttc 2382
Db 34492 CCTCCGAGCAGAGTACTCAGTGTGCCCCATCCACGTTGTCCCTGTGCGGCCCCACTTTC 34433
QY 2383 cccggcttaagcagctcctcgtgacctaaagacatcctgagcagggccagccacccatg 2442
Db 34432 CCCGGCTTAAGCAGCTGCTCGCTGAGCTTAAAGACATCTCTGAGCAGGCGCAGCACCATG 34373
QY 2443 caactccttcagcaacagcagcaggaagcagtatccgaagcagcagcagcagcagcagcag 2502
Db 34372 CAACTCCTTCAGCAACAGCAGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGCG 34313
QY 2503 agggagcctcgggggtcccaacttcccaaggaacaattacttaggggcacttcgacc 2562
Db 34312 AGGAGGCCCTCGGGGGCTCCCACTTCTCCCAAGSACAAATTAATTAGGGGCATTCGACC 34253
QY 2563 atttctgacaacgccaaggaggtgtgtaaggcagtgctcgtgtccatgggacctgggtgtg 2622
Db 34252 ATTTCTGACAACGCCCAAGGAGTTGTGTAAGGCAGCTGTGGGTGCCATGGGCTGGGTGTG 34193
QY 2623 gagcgcttgagcatctgagtcaggggaagcagcttcgggggattgcatgtacgccccca 2682
Db 34192 GAGCGTGTGAGCATCTGATCTCAGGGGAACAGCTTCGGGGGATTGCAATGTACGGCCCCA 34133
QY 2683 ctttgggagttccaccgcgtgtgctgtccactccttctgtcccccattggccgaatgcaaa 2742
Db 34132 CTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCTTGTGCCCATTTGCCGAATGCCAA 34073
QY 2743 gttctctgttagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2802
Db 34072 GGTGTCTGTGTAGCAGCAGCGGAGGCGAGAGCAGCTGAAGATATCTGTGATTTCCCT 34013
QY 2803 ttcaaggaggtttacaccaaagggttagaaggcgagcagcagcagcagcagcagcagcagcag 2862
Db 34012 TTCAAGGGAGGTTACACAAAGGCTTAGAGGCGAGAGCCTAGGCTGTCTGCGACCGCT 33953
QY 2863 gcagcagggagctccgggacacttgactgcctgtacctgtctctacagtcgga 2922
Db 33952 GCAGCAGGGAGCTCCGGGACACTTGAACCTGCCGTCTACCTGTCTCTCTACAAAGTCCGGA 33893
```

```
QY 2923 gcaactgagcagcagcagctcgctaccagagctcgactactacaacttttccactggtctcg 2982
Db 33892 GCACCTGAGAGGACGACGCTCGGTACAGAGTCCGGACTACTACTAACTTCCACTTGCTCG 33833
QY 2983 gccggaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3042
Db 33832 GCCGGACCGCGCCCTCCGCGGCTCCCATCCCGCTCCCATCCCGCTCGCATCAAGCTGGAGAAC 33773
QY 3043 ccgctgactacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3102
Db 33772 CGGCTGACTACGGCAGCGCTGGGCGCTGCGGCGGCGCAGTGCCTATATGGGACCTG 33713
QY 3103 ggagcctcagcggcgcggtgtagcgggaacccggttctgtggtaccctcagcccgct 3162
Db 33712 GCGAGCTCATGGCGGGTGCAGCGGGACCCGGTTCTGGGTACCTCAGCCGCCCT 33653
QY 3163 tctcatctctggcacactctcttcacagccgaggaagccaggtgtatggacctgtgtgt 3222
Db 33652 TCTCATCTCTGGCACACTCTCTTACAGCCGAAGAGGCCAGTTGTATGGACCGTGTGT 33593
QY 3223 ggtgtgggggtggtggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 3282
Db 33592 GGTGTTGGGGGTGGT---GGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 33536
QY 3283 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 3342
Db 33535 GCGGCGGCGGAGCGGGAGCTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTG 33476
QY 3343 gggggccaggaaagcagcttcaccgcacctgatgtgtgtaccttgcgcgcatggtgagc 3402
Db 33475 GGGGCGCAGAAAGGAGCTTACCACGACTGATGTGTGTACCTTGGCGCATGGTGAGC 33416
QY 3403 agagtccctatcccagtcacctgtatcaaacgaaatgggccccctgtagatgagc 3462
Db 33415 AGAGTCCCTATCCCGAGTCCCACTTGTGTCAAAAGCGAAATGGGCCCTGTGATGGATAGC 33356
QY 3463 tactccggaccttacggggacatcgatt 3490
Db 33355 TACTCCGGACCTTACGGGGACATGCGGT 33328
```

RESULT 7

HUMARX LOCUS	HUMARX	3231 bp	mRNA	PRI	31-OCT-1994
DEFINITION	Human androgen-receptor mRNA, complete cds.				
ACCESSION	M34233				
VERSION	M34233.1 GI:179033				
KEYWORDS	androgen receptor.				
SOURCE	Human testis and prostate cancer cell line LNCaP, cdna to mRNA.				
ORGANISM	Homo sapiens				
REFERENCE	1. (bases 1 to 3231)				
AUTHORS	Govindan,M.V.				
TITLE	Specific region in hormone binding domain is essential for hormone binding and trans-activation by human androgen receptor				
JOURNAL	Mol. Endocrinol. 4 (3), 417-427 (1990)				
MEDLINE	90258935				
FEATURES	Location/Qualifiers				
source	1..3231				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/map="Xq11.2-q12"				
	109..2829				
	/gene="AR"				
	109..2829				
	/gene="AR"				
	/note="androgen-receptor"				
	/codon_start=1				
	/db_xref="GDB:G00-120-556"				
	/protein_id="AAA51780.1"				
	/db_xref="GI:179034"				
	/translation="MEVQJGLGRVYPRPSPKTYRGAFQNLFQSVREVIQNPGRHPPEA				

QY 3616 gtcttcttcaaaagagccgtgaagggaacacagaaglacctgtgcgcagcagaatgat 3675
Db 1810 GTCCTTCTCAAAAGAGCCGCTGAAGGGAACAGAAAGTACCTCTCGCCGACAGAAATGAT 1869
QY 3676 tgcactattgataaatccgaaggaataattgtccattctgtcgtctcttggaatgttat 3735
Db 1870 TGCACATTGATAATTCGGAAGGAATAATGTCCATCTTGTCTCTCGGAATGTAT 1929
QY 3736 gaagcaggatgactctgggagccgggaagctgaagaaacttgtaactgaactacag 3795
Db 1930 GAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAACTTGGTAATCTGAAACTACAG 1989
QY 3796 gaggaagagaggtcttcagcaccacagccccactcagagacacacccccagaagctga 3855
Db 1990 GAGGAAGGAGAGGCTTCAGACACCACAGCCCCACTGAGGAGACACCCAGAGCTGACA 2049
QY 3856 gtgtcacacattgaaggctatgaatgtcagcccatcttcttgaatgtcctggagccatt 3915
Db 2050 GTGTACACATTGAAGGCTATGAATGTACGCCATCTTCTGAAATGTCTTGGAGCCATT 2109
QY 3916 gagcaggtgagtgtgtgtgacacagacacaaacacagcccgactcctttgagcccttg 3975
Db 2110 GAGCCAGGTGAGTGTGTGTGTGGACAGACAAACACAGCCCGACTCTTTTGAGCCCTG 2169
QY 3976 ctctctagcctcaatgaactgggagagagacagctgttacacgtggtcaagtgggccaag 4035
Db 2170 CTCCTAGCCCTCAATGAAGTGGGAGAGACAGACAGCTGTACAGTGTGCAAGTGGGCCAAG 2229
QY 4036 gcttgcctggcttcgcgaacttacacgtggaagaccagatggtgtcattcagttactcc 4095
Db 2230 GCTTGGCTGGCCCTCCGCAACTTACAGTGGACAGACACAGATGGGTGTCATTAGTACTCC 2289
QY 4096 tggatggggtccatggtgttgcctggcgtggcgtcccttcccaatgtcaactccagg 4155
Db 2290 TGGATGGGCTCATGTGTGTGCTGCGATGGGTGGGATCTTACCAGTGTCAACTCCAGG 2349
QY 4156 atgtctacttcgcccctgatctgtgttttcaatgagtcacgcgatgcacaaagtcggatg 4215
Db 2350 ATGCTCTACTTGGCCCTGTGCTGTGTTTCAATGAGTACCGCATGCACAAAGTCCCGGATG 2409
QY 4216 tacagccaggtgtcccaatgagcagcactctctcagaggttggatggctccaaatcacc 4275
Db 2410 TACAGCCAGTGTGTCCGAATGAGCGACTCTCAAGAGTTTGGATGGCTCCAAATCACC 2469
QY 4276 cccaggaattcctgtgcataagcactgctactcttccagcattatccagtggtggg 4335
Db 2470 CCCAGGAATTCCTGTGCATGAAGCCATGCTACTCTTACGATTTATTCAGTGGATGGG 2529
QY 4336 ctgaaaaatacaaaattcttctgatgaacttcgaatgaactacatcaaggaactcgatcgt 4395
Db 2530 CTGAAAAATCAAAAATTTCTTGTATGAATTCGAATGAATGAACTACATCAAGGAATCGATCGT 2589
QY 4396 atcattgatgcaaaagaaaaatccacatcctgtcctcaagagcttccaccagctcacc 4455
Db 2590 ATCATTGCAATGCAAAAGAAAAATCCACATCTGCTCAAGAGCTTTCACAGCTCACC 2649
QY 4456 aagctcctggactccgtgcagcatttgcgagagagctgcatacgttccacttttgacctg 4515
Db 2650 AAGCTCTGGACTCCGTGCAGCTATTTCGAGAGAGCTGCATCAGTTCATCTTTGACCTG 2709
QY 4516 ctaatcaagtcacacatggtgaagcgtggacttccgggaataatgatggcagagatcatctct 4575
Db 2710 CTATCAAGTCAACATGTTGAGCGGTGGACTTTCCGGGAATGATGCGACAGATCATCTCT 2769
QY 4576 gtgcaagtgcccaagatcctttctgggaagtcaagcccatctatttccacaccagtgga 4635
Db 2770 GTCAAGTGCCCAAGATCCTTTCTGGAAAGTCAACCCCATCTATTTCACACCCAGTGA 2829
QY 4636 agcattggaacccctatttccccccccccagctcgtatgccccccttttagatgttcttgcct 4695
Db 2830 AGCATTGGAAACCCCTATTTCCTCCACCCAGCTCATGCCCTTTTCAGATGTCTTCTGCTCT 2889

QY 4696 gttataactctgcactaccctctcgtcagtgcccttgggggaatttctctatttctgtatgaag 4755
Db 2890 GTTATAACTCTGCACCTACCTCTCTGCAGTGCCTTGGGAATTTCTCTATTGATGATACAG 2949
QY 4756 tctgtcatgaacatgttctcgaattctatttgcctggccttttttttctctctctctctct 4815
Db 2950 TCTGTCATGAACATGTCTCTGAATCTATCTGCTGGGCTTTTTTTTCTCTCTCTCTCT 3009
QY 4816 ttcttttcttctcctccctcctatctaacctcccatgacacacacacacacacacacac 4875
Db 3010 TTCTTTTCT 3069
QY 4876 atttggtcctcctatctgtgttttgatggtgtgtgtatgcctttaaactctgtatgaatc 4935
Db 3070 ATTGTGGCT 3128
QY 4936 catatggccacgtgtcaagtgtgtgtgttttacagcactactctgtccagccacacaaa 4995
Db 3129 CATATGGCCACGTGTCAAGT 3187
QY 4996 cgtttacttcttcttctgtgcacacaggaagtt 5024
Db 3188 CGCTTTACTTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 3216

RESULT 8
HUMANRE 3210 bp mRNA PRI 31-OCT-1994
LOCUS Human androgen receptor mutant gene, mRNA, complete cds.
DEFINITION M73069.1 GI:178655
ACCESSION M73069.1
VERSION 1
KEYWORDS androgen receptor.
SOURCE Homo sapiens Testis cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3210)
AUTHORS Govindan, M.V.
TITLE Specific region in hormone binding domain is essential for hormone binding and trans-activation by human androgen receptor
JOURNAL Mol. Endocrinol. 4 (3), 417-427 (1990)
MEDLINE 90258935
FEATURES
Location/Qualifiers
1..3210
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="Testis"
/map="Xq11.2-q12"
616..2808
/gene="AR"
616..2808
/gene="AR"
/note="mutant gene"
/codon_start=1
/db_xref="GDB:G00-120-556"
/product="androgen receptor"
/protein_id="AAA51735.1"
/db_xref="GI:178655"
/translation="MQLQQQQQEAHVSESSSSRRAREASGAPTSKDNLTGGTSTISD
NAKELKAVSVSMGLGVAELEHUSPGEURGDCMYAPLIGVPPAVRPTFCAPLAECKG
SLDDDSAGKSTDAEYSPFKGYTKLESGSCGSAAGSSGSLFELPSTLSLYKS
GALDEAAAYQSRDYINFLALAGPPPPPHPHARIKLENLIDYGSAAWAAACORRY
GLDASLHAGACGAGPGSGPSAAASSSSHTLFTAEEQOLYPCGCGGGGGGGGGGG
EGAVAPYGYTRPPQGLAGSDFTAPDVWYPGWVSRVPYSPCTCKKSEMGPMDSY
SGVPGMRLETAARDHVLPIIDYFPPKTCILGDEASGCHYGALTCGSKVFKRAAE
CKOKYLCSNRDCTIBKFRKNCPSCRLKRCYEAGMTGLARKLKLGNLKLQEEGAS
STTSPEETTKLTVSHIEGECQIFLNLVLEAIEPVGVCAGHNNQPDSPFAALLSYL
NELGEQQLVHVWAKALGLRLNLDVHVDQMAVIQYSWGLMVFAMGKMLFTFVNSRL
YFAPDLVFNERYRMHKSMTQCVRMHLSQEFGLQITPOEFLCMKAMLIFFPELLLF
SLIPDGLKNQKFFDELRLMYIKELDRIIACKRNPTSCSRFYOLTKLLDSVQPLAR
ELHQFTFDLLIKSHVSVDEPEMWAELISVQVPKILSGKVKPIYFHTQ"

BASE COUNT 703 a 930 c 875 g 702 t
ORIGIN

	Query Match	59.4%;	Score 3019.2;	DB 9;	Length 3210;
	Best Local Similarity	97.4%;	Pred. No. 0;		
	Matches 3145;	Conservative	0;	Mismatches	23; Indels
					62; Gaps
1816	Q	gggaggcggggtaagggaagtagtggaaagattcagccaagctcaagatcgaaagtgcag	1875		
7	D	GGGAGCGGGGTAAAGGAAGTAGTGGAAATTCAGCCAAGCTCAAGATGGAAGTGCAG	66		
1876	Q	ttagggctgggaagggtctaccctcgccgcgcgtcaagacctaccgaggagctttccag	1935		
67	D	TTAGGGCTGGGAAGGGTCTACCTTCGGCCGCGCTCAACACCTACCGAGAGCTTTCAG	126		
1936	Q	aatctgttccagagctgcgcgaagtgtatccadaaacccgggccccaggcacccagagcc	1995		
127	D	AATCTGTTCCAGAGCGTCGCGAAGTGATCCAGAACCCGGGCCCCAGGACCCAGAGCC	186		
1996	Q	gcgagcgagcaacctcccggcgccagcttctgtctctgcagcagcagcagcagcagcag	2055		
187	D	GCGAGCGCAGACACCTCCCGCGCCAGTTTCTCTCTCTCT-----GCAGAG	231		
2056	Q	cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagaagact	2115		
232	D	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGACT	291		
2116	Q	agccccagggcagcagcagcagcagcagcagggtgagagatggttctcccaagcccatcgtaga	2175		
292	D	AGCCCCAGGCAGCAGCAGCAGCAGCAGCGGTGAGGTGTTCTCCCCAAGCCCATCGTAGA	351		
2176	Q	ggccccagagctacctggtctctgattgaggaacagcaaaccttcacagcccgagctcgcc	2235		
352	D	GGCCCCACAGGCTACCTGGTCTCGATTGAGAAACAGCACTTTCACAGCCGAGTCGGGCC	411		
2236	Q	ctggagtcgcccccgagagaggttgcgtcccagagcgtcggagccgcgtggccgcagcgc	2295		
412	D	CTGGAGTGCACCCCGAGAGAGGTTGGTCCCAGAGCCTGGAGCCGCGTGGCGGCCAGC	471		
2296	Q	aaggggctgcgcgagcagctcccgagccctccggagcagagatgaactcagctgccccatcc	2355		
472	D	AAGGGGCTGCCGAGCTGCCAGCACTCCCGAGCAGGAGTCACTCAGTGCCCCATCC	531		
2356	Q	acgtgtccctctgtaggcccccaatttcccggcttaagcagctgctccgctgaccttaaa	2415		
532	D	AGTTGTCCCTCTGGCCCCCACATTCCCGGCTTTAGACACTGTCTCCGCTGACCTTAAA	591		
2416	Q	gacatctgagcagggccagcaccatgcaactctctcagcaacagcagcagcagcagcagta	2475		
592	D	GACATCTTGAGCGAGCCAGCACCATGCACTCTTCAGCAACAGCAGCAGGAAGCAGTA	651		
2476	Q	tcgaaaggcagcagcagcgggagcagcaggagccttcgggggctcccattctctccaag	2535		
652	D	TCGAAAGCAGCAGCAGCGGAGCAGAGGAGCGCTCGGGGGCTCCCATTTCTCTCAAG	711		
2536	Q	gacaattacttagggggcacttcgacccattctgacaacgcgaaggaggttgttaaggca	2595		
712	D	GACAATTACTTAGGGGGCACTTCGACCACTTCTGACAACGCCAAGGAGTTGTGTAGGCCA	771		
2596	Q	gtctcggtgtccatggcctgggtgltggaggcgttgagcatctgagtcagcagcagcagcag	2655		
772	D	GTCTCGGTGTCCATGGGCTGGGTCTGGAGGCGTTGGAGCATCTGTAGTCCAGGGGAACAG	831		
2656	Q	cttcggggggattgcatgtaagcccccaatttgggagtttcaaccctgtagctccact	2715		
832	D	CTTCGGGGGGATTGCATGTACGCCCCACATTGTGGGAGTTCCACCCGCTGTGCTGCCACT	891		
2716	Q	ccttgtgccccattggccgaaatgcgaaggttctctgctagacgacagcgcagcgcagaagc	2775		
892	D	CCTGTGCCCATTTGCCCGCAATGCAAGGTTCTCTCTGTAGCAGACGCGCAGGCAAGAGC	951		
2776	Q	actgaagatactgctgagtattcccccttccaaaggagaggtatcacaccaagggctagaaggc	2835		
952	D	ACTGAAGATACTGCTGAGTATTCCGCTTTCAAGGGAGGTTACACCAAGGGCTAGAAGC	1011		

QY	2836	gagagctagggtgctctgagcgcgtgcagcagggagctccggacaacttgactgcgcg	2895
DB	1012	gagagctagggtgctctggcagcgtgcagcagggagctccggacaacttgactgcgcg	1071
QY	2896	tctaccctgctctctacaagtcggagcactggacdaggcagctgcgtaccagagtcgc	2955
DB	1072	tccaccctgctctcttacaagttccgagcacttggacdaggcagctgcgtaccagagtcgc	1131
QY	2956	gactactacaactttccactggctctggccggagccgcgccccctccgcgcctccccat	3015
DB	1132	gactactacaactttccactggctctggccggagccgcgccccctccgcgcctccccat	1191
QY	3016	cccacagctcgacataagctgagaaaccgctggactacgcagcgcctggcggtgcg	3075
DB	1192	cccacacgctcgacataagctgagaaaccctggctggactacgcagcgcctggcggtgcg	1251
QY	3076	ggcgcgagtcgcgctatggagacctggcgagccttgcatggcgggtgagcgggagccc	3135
DB	1252	ggcgcgagtcgcgctatgggagaccttggcgagccttgcatggcgggtgagcgggagccc	1311
QY	3136	ggttctgggtcaacctgaagccgcgtctctcatctcctggcacactctctcaagccgaa	3195
DB	1312	ggttctgggtcaacctgaagccgcgtctctcatctcctggcacactctctcaagccgaa	1371
QY	3196	gaagggcagttgtatggaaccttggtgtgtgggggtgggtggcgcgcgcgcggc	3255
DB	1372	gaagggcagttgtatggaaccttggtgtgtgggggtgggtggcgcgcgcgcggc	1413
QY	3256	ggcg	3315
DB	1414	-----ggcg	1467
QY	3316	ggctacactcgccccctcaggggctggcgggcctcagggagaaagcgactccacgcacctgat	3375
DB	1468	ggctacactcgccccctcaggggctggcgggcctcagggagaaagcgactccacgcacctgat	1527
QY	3376	gtgtgtgtaccctggcgcatgggtgagcagatggcctatcccagtcgccacttgtcaaa	3435
DB	1528	gtgtgtgtaccctggcgcatgggtgagcagatggcctatcccagtcgccacttgtcaaa	1587
QY	3436	agcgaatacgccccctcagatggatagctactcggacactcagggagacatcggttggag	3495
DB	1588	agcgaatacgccccctcagatggatagctactcggacactcagggagacatcggttggag	1647
QY	3496	actgcgaaggacatgttttgccactgactatattctcccccagagaagacctgcctg	3555
DB	1648	actgcgaaggacatgttttgccactgactatattctcccccagagaagacctgcctg	1707
QY	3556	atctgtggagatgaagctctctgggtgtcaactatggagctctcacatgtggaagctgcaag	3615
DB	1708	atctgtggagatgaagctctctgggtgtcaactatggagctctcacatgtggaagctgcaag	1767
QY	3616	gtctcttcaaaaagcgcgtgaagggaacacagaagtacctgtggccagcagaaatgat	3675
DB	1768	gtctcttcaaaaagcgcgtgaagggaacacagaagtacctgtggccagcagaaatgat	1827
QY	3676	tgcaattatgataaattccgaaggaataattgtccatttgttgttgttgttgttgttgt	3735
DB	1828	tgcaattatgataaattccgaaggaataattgtccatttgttgttgttgttgttgttgt	1887
QY	3736	gaagcagggatgactctgggagccggagaactgaagaaacttgtaactcgaactcacag	3795
DB	1888	gaagcagggatgactctgggagccggagaactgaagaaacttgtaactcgaactcacag	1947
QY	3796	gagggaaggaggtctccagcaccacagccccactgaggagacacccagagctgaca	3855
DB	1948	gagggaaggaggtctccagcaccacagccccactgaggagacacccagagctgaca	2007
QY	3856	gtgtcacacattgaaggctatgaatgtcagcccaattcttctgaatgtcctgaagccatt	3915
DB	2008	gtgtcacacattgaaggctatgaatgtcagcccaattcttctgaatgtcctgaagccatt	2067

[illegible]

[illegible][illegible]

QY	4384	gaactgcatgctatcattgcatgcaaaagaaaaatccccacatctctgctccaagcgttc	4443
Db	2485	GAATCTCGATCTATCATTTGTCATGCAAAAGAAAAAATCCACATCTCTGCTCAAGACGCTTC	2544
QY	4444	taccagctccaccaagctccttggaactccgtgcagcctattgagagagagctgcatttc	4503
Db	2545	TACCAGCTCACCAAGCTCCTTGGACTCCGTGACGCTATTTCGAGAGAGCTGCATCAGTTC	2604
QY	4504	actttgacctgcttaatacaagtcacacatggtgagcgtgagctttccggaataatgaca	4563
Db	2605	ACTTTTGACCTGCTTAATCAAGTCACACATGGTGAGCGCTGACATTTCCGGAATGAIGCA	2664
QY	4564	gagatcatctctgtgcaagtgcccaagatcctcttctgggaagtcacaaagccatctatttc	4623
Db	2665	GAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGAAAAAGTCAAGCCCATCTATTTC	2724
QY	4624	cacaccagtgaaagcattgaaa-cctatttcccccaccagctcatgcccccttcag	4682
Db	2725	CACACCCAGTCAGCATTTGGAAACCCTATTTCCTCCACCCAGCTCATGCCCCCTTTTCAG	2784
QY	4683	atgtctctgacctgttataactctgcaactactcctctctgcagtgcccttgggaattctctc	4742
Db	2785	ATGTCTTCTGCTGTTATAACTCTGCACCTACTCCCTCTGCAGTGCCCTGGGGAATTTCTC	2844
QY	4743	tattgatgtacagtctgctatgaac	4767
Db	2845	TATTGATGTACAGCTGTCTGAAC	2869
RESULT	11		
AF162704			
LOCUS	2827 bp	mRNA	29-JUL-1999
DEFINITION	Homo sapiens cell-line MDA-MB-453 androgen receptor mRNA, complete cds.		
ACCESSION	AF162704		
VERSION	AF162704.1	GI:5639998	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Jin,C.H., Urcan-Bisel,M.S. and Schrader,W.T.		
TITLE	Androgen receptor sequences in human mammary carcinoma MDA-MB-453 cells		
JOURNAL	Unpublished		
REFERENCE	2. (bases 1 to 2827)		
AUTHORS	Jin,C.H., Urcan-Bisel,M.S. and Schrader,W.T.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUN-1999) Endocrine Research, Ligand Pharmaceuticals Inc, 10255 Science Center Drive, San Diego, CA 92121, USA		
FEATURES	Location/Qualifiers		
Source	1..2827		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/cell_line="MDA-MB-453"		
	/tissue_type="mammary carcinoma"		
	4..2724		
	/translation="MEVQLGLGRVYPRPSTKTYRCAFONLFQSVREVTONPGPRHPDAASAAPGASILLQQQQQQQQQQQQQQQEQEISPRQQQQQGGEDSPQARRRGPPTGLVLDPEQSPQPSQSALECHPQGVPEPAAVAASKGLPOOLPADPEDDSDAAPTSLYLPPTFGSLSDKDLIDLSAETMOLLQOQQQEAQSEVSGSARCAGASPTSKDNYLGSTTSSDNALCKRAVSMGLGVEALEHLSPQQLRGDCMTAPLILGVPVAVRPTPLCAPLCKSLDSDAGSTEDTAYSPFKGTYTKGLESLGCSGSAAGASCTLFLPSTLSLYKSGALDEAAQSYRYNFPILALAGPPPPPPHPHARIKLENPLOYGSNAAAAAACRYGDILASLHGCAAGPGSGSPSAASSSWHTLFTAEQCILYGPCGGGGGG		
CDS	/note="steroid hormone receptor; nuclear receptor superfamily"		
	/codon_start=1		
	/product="androgen receptor"		
	/protein_id="AAB45921.1"		
	/db_xref="GI:5639998"		

[illegible]

|||||
Db 814 GCTGTGCGTCCACACCTCTGTGCCCCATGGCCGAATGCAAAAGGTTCCTCTGCTAGACGAC 873
QY 2761 aqcgaggaagacactgaagataactactagatattccctttcaaggaggttacacc 2820
Db 874 ACCGAGGCAAGAGCACATGAAGATACTGCTGAGTATTCCTCTTCAAGGGAGGTATACCC 933
QY 2821 aaagggttagaaggcgagagcctaggtgctctgagcagctcgagcagggagagctccggg 2880
Db 934 AAAGGCTAGAAAGCGAGAGCCCTAGGCTGCTTGCGCAGCGCTCGAGCAGGAGCTCCGGG 993
QY 2881 acactgaactcgcctctacccctgtctctctacaaagtccggagcactgagcaagggcagct 2940
Db 994 ACACATGAACCTGCCGTCTACCCGTCTCTCTACAAAGTCCGGAGCACTGGAGGAGCAAGCT 1053
QY 2941 gcgtaacagagtcgcgactaactacaactttccactggctctgcccggagccgcgcacct 3000
Db 1054 GCGTACCAGAGTCGGCACTACTACAACTTCCACATGGCTCTGGCCGGACCGCGCCGCCCT 1113
QY 3001 cgcgcgctcccccatacccaacgctcgatcaagctggagaaacccgcgtggaactacaggaagc 3060
Db 1114 CCGCCGCTCCCATCCCAACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGC 1173
QY 3061 gectggcgagctgcggcgagcgagctaccactatgggacctggcgagcctgcgagctggcg 3120
Db 1174 GCTGGCGGCGCTGCGGCGGCGCAGTGCCTATGGGACCTGGCGAGCCTGCAATGGCGCG 1233
QY 3121 ggtgagcgggaccgggttctgggtcacccctcagcgccgcgttccctcatctctggcacact 3180
Db 1234 GGTGACGGGACCCGGTCTGGTTCACCTCAGCGCGCGCTTCCTCATCCTGGCACACT 1293
QY 3181 ctcttcacggccgaagaaagccagttgtatgacccgtgtggtgggtgggtggtggc 3240
Db 1294 CTCTTCACAGCCGAAGAAGCCAGTGTATGACCGGTGTGGTGGTGGGGGTGGT --- 1350
QY 3241 ggcggcg 3300
Db 1351 -----GGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1389
QY 3301 gctgtagcccccctacggtctacactcgccccctcaggggctggcgggccagaaagcgagc 3360
Db 1390 GCTGTAGCCCCCTACGGCTACATCGCGCCCCCTCAGGGGCTGGCGGCGCAGGAAAGCGAC 1449
QY 3361 ttcacgcacactgattgtgtaccctcgagcactgtagcagagagtcacctatccagat 3420
Db 1450 TTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCACT 1509
QY 3421 cccacttgttcaaaaagcgaaaatgggcccctggatggatagctactcccggaaccttacggg 3480
Db 1510 CCCACTTGTCTAAAAGCGAAATGGCCCCCTGGATGGATGCTACCTCGGACCTTACGGG 1569
QY 3481 gacatgcgtttgagagactgcagggaccatggttttcccattgaactattactttccaccc 3540
Db 1570 GACATGCGTTTGGAGACTGCGCAGGACCACATGTTTGCCATTGACTATTACTTTCCACCC 1629
QY 3541 cagaagacctgcctgactgtgaagatgaagcttcgggtgtcactatgagagctctcac 3600
Db 1630 CAGAAGACCTGCGCTGATCTGTGGAGATGAAGCTTCGGGTGTCATATGGAGCTCTCAC 1689
QY 3601 tgtggaagctgcaagggtctcttcaaaagagccgctgaagggaacacagaagtagacctgtgc 3660
Db 1690 TGTGGAAGCTGCAAGGCTTCTTCAAAGAGCCGCTGAAGGGAAACAGAAAGTACCTGTGC 1749
QY 3661 gccagcgaataattgactactattgataattccgaaggaaaaattgtccattcttgc 3720
Db 1750 GCCAGCAAAATGATTGCATATTGATTAATTCGAGAGGAAAAATGTCCATCTGTGCT 1809
QY 3721 ctccgaaattgttatgaagcagggatgactctggagcccggaagctgaagaaacttgg 3780
Db 1810 CTTCCGAAATGTTATGAAGCAGGGATGACTTGGGAGCCCGGAGCTGAAGAACTTGGT 1869
QY 3781 aatctgaactacagggagggaggggcttccagcaccaccagcccccactgaggagaca 3840
|||||

Db 1870 AATCTCAAACTACAGAGGAAGGAGAGGGCTTCAGCACACACGCCCTACTGAGGAGACA 1929
QY 3841 acccagaagctgacagtgtcacacattgaagctatgaatgtcagcccatcttctbgaat 3900
Db 1930 ACCCAGAAGCTGACAGTGTACACATTGAAGGCTATGAATGTACGCCCATCTTCTTGAAT 1989
QY 3901 gtccctggaagcatttagccaggtgtagtgtgtgtggtggtggtggtggtggtggtggtggt 3960
Db 1990 GTCTGGAAGCCCAATTGAGCCAGGTGTAGTGTGTCTGGACACGACAAACACGACCCGAC 2049
QY 3961 tcccttgagccttgcctctctagccctcaatgaactgggagagagacagcttgttacacgtg 4020
Db 2050 TCCTTTGCACGCTTGTCTCTAGCCTCAATGAACCTGGGAGAGACAGCTGTGTACACGTG 2109
QY 4021 gtaagtggccaaagcgcttgcctgccttcgcgcaacttacacgtggagcaccagatggct 4080
Db 2110 GTCAAGTGGGCCAAGGCTTGCCTGGCCCTCCGCAACTTACACGTGGACGACCAAGATGGCT 2169
QY 4081 gtcattcagtaectcctggtggtggctcatggtgttggcatgggtgggtggtggtggtggtggt 4140
Db 2170 GTCAATCAGTACTCTCTGGATGGGCTCATGCTGTTCCTATGGCTGGCGATCCTTCACC 2229
QY 4141 aatglcaactccaggaatgctctacttgcctccctgcgtggttttcaatgaagtagtaccgctg 4200
Db 2230 AATGTCAACTCCAGGATGCTCTACTTCGCCCTTGATCTGCTTTCAATGAGTAGTACCGCATG 2289
QY 4201 cacaagtcceggatgtacagccaggtgtccgaataagcaccctctcctcaagagtttggga 4260
Db 2290 CACAAGTCCCGGATGTACGCCAGTGTCCGAATGAGGACCTCTCTCAAGAGTTTGGGA 2349
QY 4261 tggctccaaatacccccaggaattcctgtgcatgaagcactgctactcttcttcagcatt 4320
Db 2350 TGGCTCCAAATCACCCCCAGGAATTCTGTGCATGAAAGCATGTCTACTCTTCAGCATT 2409
QY 4321 attccagtgagtggtgtgaaataacaaaattcttctgactgaactcctcaatgaactacac 4380
Db 2410 ATTCAGTGGATGGGCTGAAATAATCAAAAATTCCTTGTGAATTCGAATTCGAATGAACTAC 2469
QY 4381 aagaaactcgtatgatactgtgcatgcaaaagaaataccacatcctgctcgaagcgc 4440
Db 2470 AAGNACTCGATGTATCATTTGCATGCAAAAGAAATAATCCACATCTCTGCTCAAGACG 2529
QY 4441 ttctaccagctcacaagctcctggaactccgtgcagcctatgtcgagagagctgcatacg 4500
Db 2530 TTCTACCAAGCTCACCAAGCTCCTTGACTCCGTGCATCCTATTGGGAGAGAGTGCATCAG 2589
QY 4501 ttcaactttgacctgttaatacaagtccacatgagcgtggacttccggaataatgag 4560
Db 2590 TTCACCTTTGACCTGCTAATCAAGTCACATGGTGAGCGTGGACTTTCGCGAAATGATG 2649
QY 4561 gcagagatcctctgtcgaagtgcgaagatccctttctgggaagtcgaagccatctat 4620
Db 2650 GCAGAGATCATCTGTGCAAGTGCCTCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTAT 2709
QY 4621 ttccacacccagtggaagcattggaaacccctatttccccacccccagctatgcccccttc 4680
Db 2710 TTCCACACCCAGTAGAAGCATTTGAAACCCCTATTTCACACCCAGCTCATGCCCTTTC 2769
QY 4681 agatgtcttctgcctgtataactctgcactactcctgcagtccttgggaattt 4738
Db 2770 AGATGTCTTCTGCTGTATATACTCTGCACACTCTCTGCGAGTCCCTTGGGAATTT 2827

RESULT 12

MFU94179
LOCUS MFU94179 2821 bp mRNA PRI 16-SEP-1999
DEFINITION Macaca fascicularis androgen receptor mRNA, complete cds.
ACCESSION U94179
VERSION U94179.1 GI:3861480
KEYWORDS :
SOURCE crab-eating macaque.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
1 (bases 1 to 2821)
Choong,C.S., Kempainen,J.A. and Wilson,E.M. Evolution of the primate androgen receptor: a structural basis for disease
J. Mol. Evol. 47 (3), 334-342 (1998)
98404153
PUBMED
9732460
2 (bases 1 to 2821)
Choong,C.S., Kempainen,J.A. and Wilson,E.M. Direct Submission Submitted (18-MAR-1997) Laboratories for Reproductive Biology, University of North Carolina at Chapel Hill, CB #500, MSRB, Rm 370, UNC-CH, Chapel Hill, NC 27599, USA
Location/Qualifiers
1. .2821
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
1. .2688
/codon_start=1
/product="androgen receptor"
/protein_id="AAC73050.1"
/db_xref="GI:3861481"
/translation="MEVQLGLGVRPPRSKTYRGAFONLFQSVREVIONPGRHPHEA ASAAPCASLQQOQQOQETSPROOOQOQEGDSPOAHRGPTGYLYLDEEOQSQP QSAPECHPERGCYEPGECAVAACGKLPOQLPAPDEDNSAPSTLSLLGFTFFGLSSC STDLKILSEASTNMLLOQQOQEAESGSSGRAEASGAPTSSKNITLGTTSTIDS AKELCKAVSMVLGVLEALBHLHPSGQLRDCMYPVLGPVPARTTQCPLAECKGS LLDDSAKSTEDYTFSPFKAGGT KGLESGSCSAAAGSSTLELPSTLSLXSG ALDEAAAYQRDYNFPLLAGTPPPPDPHPHARIKLENPLDYGSAAWAAAQAQRYG DLASHGAGAAGPGSPSAAASSWHTLFTAEGOLYGPCGGGGGGGGGGAGGA GAVAPYCTRPPOGLAGEGDFTAPDVWYPGMVSRVPVSPCTVKSEMGPMWDSYSG PYGMRLTEARDHVLPIDYVFPOKTCLCGDEASGCHGALTCGSKCVFFKKRAAGIK QKYLCSARNDCITDKPRKNCPSCLRKLCEAGMTIGARKLKLNKLQOEAGEAST TSPTETPAQLTVSHTEGECQPIFLNVLEIEPGVVCGHDNNPDFAALLSSLNE LGBRELPHVHWAKALPGRNLHVDQDMAVIQYSWMGLMYFMWGSRFSFTNVMSRLFY ADPLEVNFEMHSRYSQCFNRHLSQEFGLWIQTPOEFLCMKALLFSIIPVDGLK NOKFDBLRNYIKELDIRIACKRNPTCSRFRYOLTLLDSVQPIARELHQFTFDL LIKSHVSDVDFPDMAEILISVQPKILSGKVKIYHTQ"
BASE COUNT 623 a 829 c 799 g 570 t
ORIGIN
Query Match 51.4%; Score 2612.4; DB 9; Length 2821;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 2771; Conservative 0; Mismatches 46; Indels 91; Gaps 4;
Qy 1864 atgaagtgcagtttagcctggaaagggtctaccccgccgctccaagaacctaccga 1923
Db 1 ATGAGGTGCAGTTAGCGCTCGGGAGGGGTCTACCCTCGCCGCCGCTCCAGACTACCGA 60
Qy 1924 ggagcttccagaatctgttccagagcgtgcgcaagtgatccagaaccgccgccccag 1983
Db 61 GGAGCTTTCAGAAATCTGTTCAGAGCGTGCGCGAAGTAGATCCAGAACCCGGCCCCCAGG 120
Qy 1984 caccagagcccgcgagcgagcacctccgcgcgcaggttgtgtgtcgtcagcagcag 2043
Db 121 CACCCAGAGCCGCGAGCGAGCACCATCCCGCGGCCAGTTTT----- 161
Qy 2044 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2103
Db 162 -----GGAGCAGCAGCAGCAGCAGCAG 180
Qy 2104 cagcaagagactagccccag ----cagcagcagcagcagcaggggtgagatggtctccc 2160
Db 181 CAGCAGAAAACTAGCCCCCGGCACACGACGACGACGACGACGAGCGGTGAGATGTTCTCCC 240
Qy 2161 caagcccatcgtagagccccacagactacctggttcctgatatgaacacagcaacctca 2220
Db 241 CAAGCCCATGTAAGAGSCCCACAGGCTACCTGGTCTCTGGATGAGGAACAGACGCTTCA 300
Qy 2221 cagcgcagtcgccttgagttgctccaccocagagaggttggttccagagcctgagcc 2280

Db 1354 GCTGTAGCCCTTACGGCTACACTCGGCCACCTCAGGGCTGGCGGCCAGGAAAGGCAC 1413
Qy 3361 ttccacgcacattgattgtgacccctggcggaatggtgagcagagtgccctatccccagt 3420
Db 1414 TTCACCGCACCTGATGTGTGATACCTTGGCGGCATGTTGAGCAGTGCCTATCCAGT 1473
Qy 3421 cccacttgctcaaaagcgaaatgggcccctggatggatgagctactccggaaccttaagg 3480
Db 1474 CCCACTTGTCTCAAAAGCGAGATGGGCCCTGGATGGATAGCTACTCCGACCTTACGGG 1533
Qy 3481 aacatgcttgagagctgacagagaccatgttttgccattgactattactttccaccc 3540
Db 1534 GACATGCGGTTGGAGACTGCCAGGACCATTGTTTGGCAATTGACTATTACTTTCCACCC 1593
Qy 3541 cagaagacgctgctgatctgtggagatgaagcttctgggtgctactatggagctctca 3600
Db 1594 CAGAAGACTGCTGATCTGTGGAGATGAAGCTTCTGGGTGTCATATGGAGCTCTCACA 1653
Qy 3601 tgtgaagctgcaaggtctctcaaaagacgcgtgaaggaaacagaaatgactatgac 3660
Db 1654 TGTGAAGCTGCAAGGTCTCTTCAAAAGAGCGCTGAAGGGAACAGAGTACTGTGT 1713
Qy 3661 gccagcaaaatgattgacattgataaattccgaagaaaattgtccatcttgcgt 3720
Db 1714 GCACAGCAAAATGATTGACATATGATAATTCGAAGAAAATTGTCCATCTTGCCT 1773
Qy 3721 ctccgaaatgttgaagcagggatgactctgggagcccggaagctgaagaaacttggt 3780
Db 1774 CTTCCGAAATGTTATGAAGCAGGGATGACTCTGGAGCCCGAAGCTGAAGAACTTGGT 1833
Qy 3781 aatctgaactacagaggaagagaggtctccagcaccaccagcccaactgagagaca 3840
Db 1834 AATCTGAAGCTACAGGAGGAAGAGAGGCTCCAGCACCCAGCCCTGAGGAGACA 1893
Qy 3841 accgaagctgacagtgcacacattgaagctatgaatgtcagcccatctttctgaat 3900
Db 1894 GCCCAGAACTGACAGTGTACACATTGAGGCTATGAATGTACGCCCATCTTCTGAAT 1953
Qy 3901 gtccctggaagccattgagccaggtgtagtgtgtgtgtgagacagacacacacagccagc 3960
Db 1954 GTCTTGGAAAGCCATTGAGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2013
Qy 3961 tcccttgagccttctctctagctcaatgaactgaactgagagagagacagctgtacacgtg 4020
Db 2014 TCCTTCGAGCCTTGTCTCTTCACTCAATGAATGAATGAGGAGAGACAGCTTGTATGTG 2073
Qy 4021 gtcaagtggccagccttgcctgcttccgcaacttaccagctgagacagatggct 4080
Db 2074 GTCAAGTGGCCAAAGCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2133
Qy 4081 gtcattoagtaactcctggatggggctcattggtttgttgcattggcctggcgtccctcacc 4140
Db 2134 GTCAATTCAGTACTCCTGGATGGGGCTCATGTTGTTTGGCATGGGCTGGCATCTTCAAC 2193
Qy 4141 aattgaactccagagctctactctgctcctgctgctgctgctgctgctgctgctgctgct 4200
Db 2194 AATGTCAACTTCCAGGATGCTCTACTTTGGCCCTGATCTGGTTTCAATGAGTACCCCATG 2253
Qy 4201 cacaagtcccgatgtacagccaggtgtgtccgaatgagggcacctctctcaagagtttga 4260
Db 2254 CACAAGTCCCGGATGTACACCCAGTGTGTCGAATGAGGCACCTCTCTCAAGAGTTTGA 2313
Qy 4261 tggctcacaatcaccctcccgaggaattcctgtgcatgaagagcactgtactcttcagcatt 4320
Db 2314 TGGCTCCAAATCAACCCCTCCAGGAATTCCTGTGATGAAGCGCTGTACTTCTTCAAGCAT 2373
Qy 4321 attcagtgatggctgaaaaatcaaaatcttctgctgaacttcgaatgaactacatc 4380
Db 2374 ATTCCAGTGGATGGCTGAAAAATCAAAATCTTTTGTGAATCTTGAATGAATACATC 2433
Qy 4381 aaggaaactcagctatcatctgcatgcaaaagaaaaatccccacatctctgctcaagacgc 4440
Db 2434 AAGGAACCTCGATCGTATCATTCATGATGATGATGATGATGATGATGATGATGATGATG 2493

Qy 4441 ttctaccagctcacaagctcctggaactccgtgagcctattgagagagctgcatcag 4500
Db 2494 TTCTACCACTCACCAGCTCCTGGACTCCGTCGACCTATTTCGAGAGAGCTGCATCAG 2553
Qy 4501 ttcaactttgacctgctaatacgaatcacacatggtgagcgttgactttccggaataatg 4560
Db 2554 TTCACTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCGGAATATG 2613
Qy 4561 gcagagatcatctctgcaagtgcacagatccttcttggaagaaagtcgaagccatctat 4620
Db 2614 GCAGAGATCATCTCTGTGAAGTGCCCAAAATCCTTTCTGGGAAGTCAAGGCCATCTAT 2673
Qy 4621 ttccacaccagtggaagcattgaaa-cctattttcccaaccccaagcctcatcccccttt 4679
Db 2674 TTCCACACCCAGTGAAGCATTTGGAATCCTATTCTCACCACCATCATGCCCTTTT 2733
Qy 4680 cagatgtctctgctgttataaactctgcactactcctcgtcagtgccctgggaatttc 4739
Db 2734 CAGATGCTCTTCTGCTTATAACTCTGCACTACTCTCTGCACTGCTTGGGAATTTC 2793
Qy 4740 ctctattgattacagctctgcatgaac 4767
Db 2794 CTCTATTGATGTACAGTCTGTGTCATGAAC 2821

RESULT 13
PHU941176
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

PHU941176 2769 bp mRNA PRI 16-SEP-1999
Papio hamadryas androgen receptor mRNA, complete cds.
U941176
U941176.1 GI:3861474
baboon.
Papio hamadryas
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Papio.
1 (bases 1 to 2769)
Choong,C.S.; Kempainen,J.A. and Wilson,E.M.
Evolution of the primate androgen receptor: a structural basis for
disease
J. Mol. Evol. 47 (3), 334-342 (1998)
98404153
9732460
2 (bases 1 to 2769)
Choong,C.S.; Kempainen,J.A. and Wilson,E.M.
Direct Submission
Submitted (18-MAR-1997) Laboratories for Reproductive Biology,
University of North Carolina at Chapel Hill, CB 7500, MSRB, Rm 370,
UNC-CH, Chapel Hill, NC 27599, USA
Location/Qualifiers
1. .2769
/organism="Papio hamadryas"
/db_xref="taxon:9557"
1. .2688
/codon_start=1
/product="androgen receptor"
/protein_id="AAC73047.1"
/translation="MEVOLGLGRVYPRPPTKTYRGAFQNLFSQVREVIQNPGRHPPEA
ASAPCHPEPCVPCAAVAAGKGLPOOLPAPDEDASAPSTLSLGLPTFPLGSSC
ASAPCGASLQQQQQQQQQQSPRQQQQQGGEDGSPQAHRRGPTGLVLDEEQSPQ
SADLKDLSEASTMOLLQOQQQAVSSGSSGRARASAPTSKKNYLGSTISDS
ALECKAVSVSMGLGVLEAHLSPFQDLRDCMYAVLGVPPAVRPTCPAPLAEEKGS
LLDDSAKSTEDTAEYSPFGKTYKGLGSLGSCGSAAGSGSTLELSTLSLYKSG
ALDAAAYQSRDYNFPFALAGPPPPPPPPHARIKLENPLDYGSAWAAAAQCRIG
ELASLHGAGAGGSGSPSAASASHTELEAEGLYCGCGGGGGGGGGGAGGEEA
GAVAPGYTRPQGLAGDEFTAPDWMVPGMVSRYVPSPTCVKSEMPGMDYSYSG
PYGDMRLTARDHVLPTDYEPQKCLICGDEASGCHYALTCGCKVFFKRAEAGK
QKYLICASRNOCITDKFRKNCPCRLKCYEAGMTLGARKLKLGKLNKLOEEGEAST
TSPTETEAQKLTVSHIEYCQIFLNVLEATPEGVVACGHNDNPDSPFANLLSSLINE
LGERQLVHVVKWAKALPFRNLHVDDQMAVLIQYSWMGLVMFAMGWSFTNWNRLIFY

APDLVFNEYMHKSRMYSCQVRNHLSSQEFGLWLOITPQEFCLMKALLFSIIPVDGLK
NQKFDRLRWNYIKELRIITACKRNITSCSRFYQLTKLSDVSQIARELHQFTEDL
LIKSHMYSVDFPPENMAEIIISVQVFKILSGVKVPYFHTQ"
612 a 815 c 788 g 554 t
BASE COUNT
ORIGIN

Query Match	50.8%	Score 2579.4	DB 9	Length 2769
Best Local Similarity	95.4%	Pred. No. 0		
Matches 2722	Conservative 0	Mismatches 46	Indels 85	Gaps 3
Qy 1864	atgdaagtgcagtttagggctgggaaggtctacacctcgccgctccaagacctaacga	1923		
Db				
1	ATGGAGTGCAGTTAGGGCTGGGAGGGTCTACCTCGCGCGCGTCCAGACCTACCGA	60		
Qy 1924	ggagctttccagaatctgtttccagagcgtgoggaagtgtccagaaccggcgcccgag	1983		
Db				
61	GGAGCTTTTCAGAAATCTGTTCAGAGCGTGGCGAAGTGTATCAGAAACCCGGGCCCGAG	120		
Qy 1984	caccacagggccgcgagcgcagcacctcccggcgccagtttgtctgtcagcagcag	2043		
Db				
121	CACCACAGAGCGCGAGCGCAGACACCTCCCGGCCCGAGTTT-----	161		
Qy 2044	cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	2103		
Db				
162	-----GCAGCAGCAGCAGCAGCAGCAG	183		
Qy 2104	cagcaagagactgcccccagcagcagcagcagcagcagcagcagcagcagcagcagcag	2163		
Db				
184	CAGCAAGAACTAGCCCCCGCGCAGCAGCAGCAGGGTGAGATGGTTCTCCGCCAA	243		
Qy 2164	gcccctgtagagccccacagcgtacctgctgctggatgaggaaacagcaaccttcacag	2223		
Db				
244	GCCATCTGTAGAGCCCCACAGGGTACTCTGCTGTGATCAGGAACAGCAGCCTTCACAG	303		
Qy 2224	ccgcagtcggccctgaagtccaccgcagagaggtttgctccagagcctggagcgcc	2283		
Db				
304	CCTCAGTCAGCCCCGAGTGCACCCGAGAGAGTTGTCTCCAGAGGTTCGCCAGAGCCTGGAGCGGCC	363		
Qy 2284	gtggccgcagcaagggtgtccgcgagcagctccagcacctccgagcagaggtgactca	2343		
Db				
364	GTGGCCGCCGGCAGAGGGCTTGC CGCAGCAGCTGCCAGCAGCCTCCGCCAGGATGACTCA	423		
Qy 2344	gttgcgccatccacgttgtccctgtcgtggccccaacttcccgggttaagcagcgtctcc	2403		
Db				
424	GCTGCCCATCCAGTTGTCTGTCTGTGGGCCCCACCTTCCCGCGCTTAAAGCAGCTGCTCC	483		
Qy 2404	gctgaacctaaagacatcctgagcaggccagcaccatgcacacctcttcagcaacagag	2463		
Db				
484	GCGGAGCTTAAAGACATCTGTAGCGAGGCCAGCACCATGCAACTCCTTCAGCANACAGCAG	543		
Qy 2464	caggaaagcagatccgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	2523		
Db				
544	CAGANAGCAGTATCCGAAGCGACGACGCGGGAGAGCGAGGAGGCCCTCGGGGGCTCCC	603		
Qy 2524	actctccaaagacaattacttaggggcacttgaccattctgacaaaccccaagag	2583		
Db				
604	ACTTCTCCAAAGGACAATTTACTTAGGGGCACTTCGACCATTTCTGACAGCCCAAGGAG	663		
Qy 2584	tgtgtaaagcagtgctcgtgtccatggtcctggtgtgtaggcgtttggagcatctgagt	2643		
Db				
664	CTGTGTAAAGGCACTGTCGCTGTCCATATGGCTTGGGTGTGGAGCGCTTGAGCATCTGAGT	723		
Qy 2644	ccaggggaaacagcttcgggggattgcattgacgcccacatttggaggttccaccgct	2703		
Db				
724	CCAGGGGAACAGCTCGAGGGGATTCATGTATGACGCCCCAGTTTTCGGGAGTTCCACCCCGCT	783		
Qy 2704	gtcgtcccaactccttgcgcccatgtggccgaatgcaaggttctctgtcgtagacagc	2763		
Db				
784	GTGCGTCCACATCCCGTGTGCCCATTTGGCCGAATGCAAAAGGTCTCTGCTAGACACAGC	843		
Qy 2764	gcaggcaagcagcactgaagatactgtagtattccctttcaaggagggttacacaaa	2823		
Db				

Db	844	GCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCTCTTTCAAGGGAGGTTACACCAA	903
QY	2824	gggctagaagcgagagcctaggctgctctggaagcgtcagcagggagcctccgggaca	288
Db	904	GGGCTAAGAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGGACA	963
QY	2884	cttgaactcgctctaccctgtctctctacaagtccggagcactgacagagcagctgcg	294
Db	964	CTTGAATCGCCTCCACCTGCTCTCTACAAGTCCGAGCACTGGACGAGCGAGCTGGG	102
QY	2944	taccagagtcgagactatacaattctccactgctcgccggaacgcgcgcctccg	300
Db	1024	TACCAGAGTCGCGACTACTACAACCTTTCACCTGCGCTCTGGCGGGCGCGCGCCCTCCA	108
QY	3004	cgcgctccccatccccacgctcgcatcaagctggagaaacccgctggactacagcagcgcc	306
Db	1084	CGCGCTCCCATCCCACTCGCATCAAGCTGGAGAACCCGCTGGACTATTGGCAGCGCC	114
QY	3064	tgggcgctgcggcgcgagtcgcgtatgggaacctggcgagcctggcgagcctgcgaggt	312
Db	1144	TGGCGGCTTGGCGGGCGCACTGCGCGCTATTGGGAACTGGCGAGCTGCATGGCGCGGT	120
QY	3124	gcagcgggacccggtcttgggtccacctcagcgcgcgcttctctatctctggcacactctc	318
Db	1204	GCAGCGGACCGCGCTCTGGGTCAACCTCAAGCGCGCGCTTCCTCATCTTGGCACACTCTC	126
QY	3184	ttcacagcgaagaagcgcaattgtatggaccctgtgtggtggtgggggtggtggcggc	324
Db	1264	TTTCACAGCCGAAGAAGGCCAGTTGTATTGGACCGTGTGGTGT-----	130
QY	3244	ggcg	330
Db	1306	-----GGGGCGCGCGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGCT	135
QY	3304	gtagccctcagcgtcacactcggccccctcaggggctggcgggcgaggaagcacttc	336
Db	1357	GTAGCCCCCTATGCTACACTCGCGCCACCTCAGGGGCTGGCGGGCGAGGAAGCGACTTC	141
QY	3364	acgcacactgatgtgtggtaccctggcgcatggtgacagagtccctatcccagctcc	342
Db	1417	ACCGCACCTGATGTGTGTTACCTTGGCGCGCATGGTGACGACAGTGCCTATCCGACGTCC	147
QY	3424	acttgtcaaaagcgaatggggccctggatggatagctactccggacctacggggac	348
Db	1477	ACTTGTGTCTAAAAGCAGATGGGCCCCCTGGATGGATAGTACTCCGGACCTTACGGGGAC	153
QY	3484	atgcgcttgagactcgcaaggacactgtttgccattgactattacttccaccccg	354
Db	1537	ATGCGTTTGAGACTGCCAGGACCAATGTTTGGCCATTGACTATTCTTTCCACCCCGAG	159
QY	3544	aagacctgcctgatctgtggagatgaagctctgggtgtcactatggagctctcacatgt	360
Db	1597	AAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGAGGCTTCACATGT	165
QY	3604	ggaagctgcaagctctcttcaaaagacgcgtgaagggaacagaagctacctgtgcgc	366
Db	1657	GGAGCTTGAAGGCTTCTTTCAAAGAGCCGCTGAAGTAAACAGAACTACCTGTGTGCC	171
QY	3664	agcgaatgatgtgactattgataaattccgaaggaaaaattgtccactctgtcgctctt	372
Db	1717	AGCAAAATCATTTGCATATTGNTAAATTCGAAGGAAAAATTGTTCATCTTGGCGTCT	177
QY	3724	cggaaatgttatgaagcaggatgactctgggagcccggaagctgaagaacttggtaat	378
Db	1777	CGGAAATGTATTGAAGCAGGGATGACTCTTGGCAGCGCCGGAAGCTGAAGAACTTGGTAA	183
QY	3784	ctgaactacaggagaagagagcttccagcaccacccacccactcaggagacaacc	384
Db	1837	CTGAACCTACAGGAGGAGAGAGGCTTCCAGCACCCACCGCCCTCTGAGGAGACAGCC	189
QY	3844	cagaagctgacagtgctcacacattgaaggcttatgaatgcagccactcttcttgaaatgc	390
Db	1897	CAGAAGCTACAGTGTACACATTGAAGGCTATGAATGTCAAGCCATCTTTCGAATGTC	195

Db	3957	GAGTTTGTGTATTTCTTCAAGTCGTGTCATGATCTCTTTGTGGCCCAAGTCGAA-CTGT	4015
Qy	4959	gcttgattacagcac--tactctgtccagcagcacacacaaacgtttacttattatgcac	5016
Db	4016	CTTTGTTTATAGCACCTCTGCTGTGTCGCCAACCAAGCAAAATGTTTACTCACCTTATGCCAT	4075
Qy	5017	gggaagtttagagagcLaagattatctgtgggaaatcaaaaaaacaagcaaacaa	5074
Db	4076	GGCAAGTTTAGAGCTATAAGTATCTTGGGAAGAAACAAACAGAGAGTAAAAAA	4133
RESULT	15		
RATANDREC			
LOCUS			
DEFINITION	Rat androgen receptor (AR) mRNA, complete cds.	27-APR-1993	
ACCESSION	M20133		
VERSION	M20133.1	GI:202895	
KEYWORDS	androgen receptor.		
SOURCE	Rat, epididymal cDNA to mRNA, clone rAREpl.		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 4137)		
AUTHORS	Tan, J., Joseph, D.R., Quarmby, V.E., Lubahn, D.B., Sar, M., French, F.S. and Wilson, E.M.		
TITLE	The rat androgen receptor: Primary structure, autoregulation of its messenger ribonucleic acid, and immunocytochemical localization of the receptor protein		
JOURNAL	Mol. Endocrinol. 2, 1276-1285 (1988)		
MEDLINE	89112209		
COMMENT	Draft entry and computer readable sequence [1] kindly submitted by E.M. Wilson, 18-AUG-1988.		
FEATURES	Location/Qualifiers		
source	1..4137		
CDS	/organism="Rattus norvegicus" /db_xref="taxon:10116" 994..3702 /note="androgen receptor" /codon_start=1 /protein_id="AAA40733.1" /db_xref="GI:202896" /translation="MEVQLGVRVPRPKSTYRGAFONLFQSVREATIONPGPRHPEA ASIPAGCLQROETSPFRRRRQHPEDDPDADHRTGTGYLALKEEQPSQSOAS EGHPESCLPEPGAATAPGKLPQPPAPDDSDAPSTLSLLGFTPEGLSSCSADI KDILSEAGTWLQQQQQQQQQQQQQQQQQSEVISEGSSVRAREATGAPSSK DSYLGNSTTSDSAKELKAVSMGLVGALEHLSPEQLDRGDMYASLLGCGPAVR PTPCALKECKGLSLDEGPKGTETAEYSFKSGYAKGLESGCSGAGSSGT LEIPSSLYKSGVDEAAVQNRDYNFPLAGPPHPPTPHPHARTKLENPSDY SAAAAAQCRRYGLDSLHSGSVAGSTGPPATASSWHLTFLTAPEGVLPGGGG WSSPDGVPAPYITRPPQGLASQEGDFSASEWYPPGVNRPVPSPCVKSEMG WMENYSGPYMRDLSDHVLPIDYFFPQKCLICGDEASCHYGALTCCSKVFE KRAAFGKQKYLCSNRDCTDKFRKNCPCRIKCYEAGTILGARKLGNLKLQE ECENSSAGSPTEPDSQKMTVSHIEGYEQPIFLNLEATEPGVCVAGHNNQPDFA LSNLNELGEBQLVHVYKWAALPGFRNLHDDQMAVIOYSWMLGVFMGMRSTFN NSBMLYFAPDLVENEYRMHKSVMYSCQVRMHLHLSQEFMLQITPQEFMLKLLFSI IPVDGLKQKFEDELNMNYIKELDRITACKKNPTSCSRREYQLTITKLLDSVQPIAREL HQFTFDLLIKHMSVDPEPMMAEIIISVQPKILSGKVKPIYHTQ"		
BASE COUNT	979 a 1149 c 1086 g 923 t		
ORIGIN			
Query Match	48.7%	Score 2473.2;	DB 10; Length 4137;
Best Local Similarity	77.8%;	Pred. No. 0;	
Matches 3406;	Conservative	0; Mismatches	666; Indels 308; Gaps 23;
Qy	758	atccccgggagccagcttgcctggagagcgggaacggtccggagcaagccagaggagga	817
Db	1	ATCCCTAGGAGCAGCCTGCTGGAGAACGAG-AGGGTCCGAGCAACCTGGAGGCTGA	59
Qy	818	ggaagccagagagggaaaaagggcccnagctagccgctccagtcgtctacagagccgaa	877
Db	60	GAGGCGATCAGAGGGGAAAGA--CTGATTTAGCCACTTACGTGCATACAGACCTTAA	117
